การพัฒนาวิธีการประเมินการปฏิบัติงานของอาจารย์มหาวิทยาลัย: กรณีศึกษามหาวิทยาลัยสองแห่ง

นางสาว จิตรเลขา ธีระจามร

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาครุศาสตรดุษฎีบัณฑิต
สาขาวิชาการวัดและประเมินผลการศึกษา ภาควิชาวิจัยการศึกษา
คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2545
ISBN 974-17-1435-1

DEVELOPMENT OF METHODOLOGIES TO EVALATE THE PERFORMANCE SCORES OF UNIVERSITY INSTRUCTORS: A CASE STUDY IN TWO UNIVERSITIES

Miss Jitlekha Teerajarmorn

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Education Measurement and Evaluation

Department of Educational Research
Faculty of Education
Chulalongkorn University

Academic year 2002
ISBN 974-17-1435-1

Thesis Title	Development of Methodologies to Evaluate the Performance Scores of University Instructors: A Case Study in Two Universities.
By	Miss Jitlekha Teerajarmorn
Department	Educational Research
Thesis Advisor	Professor Dr.Utumporn Jamornmann
Thesis Co-advisor	Professor Dr.Peter Rowlinson
	Associate Professor Dr.Arunee Onsawad

Accepted by the Faculty of Education, Chulalongkorn University in Partial Fulfillment of the Requirements for the Doctoral Degree

..... Dean of the Faculty of Education

(Associate Professor Dr.Paitoon Sinlarat)

THESIS COMMITTEE

..... Chairman

(Associate Professor Dr.Sirichai Kanjanawasee)

..... Thesis Advisor

(Professor Dr.Utumporn Jamornmann)

..... Thesis Co-Advisor

(Professor Dr.Peter Rowlinson)

..... Member

(Dr.Chaub Leechor)

..... Member

(Associate Professor Dr.Siridej Sujiva)

จิตรเลขา ธีระจามร: การพัฒนาวิธีการประเมินการปฏิบัติงานของอาจารย์มหาวิทยาลัย: กรณีศึกษามหาวิทยาลัยสองแห่ง (DEVELOPMENT OF METHODOLOGIES TO EVALUATE THE PERFORMANCE SCORES OF UNIVERSITY INSTRUCTORS: A CASE STUDY IN TWO UNIVERSITIES) อ. ที่ปรึกษา: ศ. ดร.อุทุมพร จามรมาน, อ. ที่ปรึกษาร่วม: ศ. ดร.ปีเตอร์ รอว์ลินซัน และ รศ. ดร.อรุณี อ่อนสวัสดิ์, 167 หน้า.
ISBN

การวิจัยครั้งนี้มีวัตถุประสงค์หลัก 2 ประการ คือ ประการแรก เพื่อพัฒนาวิธีการวัดค่าการปฏิบัติงานของอาจารย์มหาวิทยาลัย ประการที่สอง เพื่อพัฒนาสูตรคำนวณค่าการเปลี่ยนแปลงการปฏิบัติงานในช่วง 3 ปี

จากการศึกษาวรรณคดีที่เกี่ยวข้องและการสัมภาษณ์ผู้บริหารและอาจารย์จากมหาวิทยาลัยธรรมศาสตร์และเซนต์จอห์น พบว่า ภาระงานของอาจารย์มหาวิทยาลัยประกอบด้วย 6 งาน รวม 35 งานย่อย ซึ่งแต่ละงานย่อยได้รับการยืนยันว่าเหมาะสมในการบ่งชี้การปฏิบัติงานของอาจารย์มหาวิทยาลัย โดยน้ำหนักของภาระงานและงานย่อยได้รับการกำหนดจากผู้บริหารของแต่ละมหาวิทยาลัย ซึ่งพบว่า น้ำหนักที่ได้แตกต่างกันตามปณิธานและภูมิหลังของแต่ละมหาวิทยาลัย

การเก็บข้อมูลการปฏิบัติงานของอาจารย์มหาวิทยาลัยธรรมศาสตร์และเซนต์จอห์น (รวม 290 คน) ในปีการศึกษา 2541, 2542 และ 2543 ข้อมูลใน 14 งานย่อย เก็บรวบรวมจากฐานข้อมูลของมหาวิทยาลัย ส่วนอีก 21 งานย่อย เก็บข้อมูลโดยใช้แบบสอบถามเพื่อรายงานการปฏิบัติงาน สำหรับการคำนวณค่าความเที่ยงโดยแอลฟาของครอนบาค (Cronbach's alpha) พบว่ามีค่าประมาณ 0.8 แสดงว่า ค่าการปฏิบัติงานที่คำนวณได้มีความเที่ยงแบบความสอดคล้องภายใน นอกจากนี้เมื่อจัดอันดับค่าความสัมพันธ์ระหว่างคะแนนงานย่อยกับคะแนนรวม แล้วหาความสัมพันธ์ระหว่างปีการศึกษาในแต่ละมหาวิทยาลัย พบว่า มีความสัมพันธ์สูง ซึ่งแสดงถึงเสถียรภาพของคะแนน จากนั้นหาความสัมพันธ์ระหว่างสองมหาวิทยาลัย พบว่า มีความสัมพันธ์ที่น้อยกว่า ซึ่งแสดงถึงความแตกต่างของความสัมพันธ์ระหว่างคะแนนงานย่อยและคะแนนรวมระหว่างสองมหาวิทยาลัย นอกจากนี้ จากการสัมภาษณ์ความคิดเห็นของหัวหน้าภาคของทั้งสองมหาวิทยาลัย เกี่ยวกับความตรงของค่าการปฏิบัติงานที่คำนวณได้ของอาจารย์ในภาค พบว่า ค่าการปฏิบัติงานของอาจารย์ร้อยละ 90 มีความตรง ดังนั้น วิธีการวัดค่าการปฏิบัติงานที่ผู้วิจัยพัฒนาขึ้น ให้ค่าการปฏิบัติงานที่มีความเที่ยงและความตรง

สูตรที่ผู้วิจัยพัฒนาขึ้นสำหรับคำนวณค่าการเปลี่ยนแปลงการปฏิบัติงาน ขึ้นกับขนาดและรูปแบบการเปลี่ยนแปลงของค่าการปฏิบัติงานในช่วง 3 ปี โดยแบ่งออกเป็น 2 สูตร คือ สูตรสำหรับการเปลี่ยนแปลงแบบเส้นตรง และสูตรสำหรับการเปลี่ยนแปลงแบบเส้นโค้ง แต่จะพบว่า อาจารย์ส่วนใหญ่ (ร้อยละ 96) มีลักษณะการเปลี่ยนแปลงแบบเส้นโค้ง ดังนั้น เฉพาะสูตรสำหรับเส้นโค้ง ได้นำมาใช้เพื่อคำนวณค่าการเปลี่ยนแปลงการปฏิบัติงานของอาจารย์แต่ละคน ซึ่งค่าที่ได้จากสูตรดังกล่าว ได้รับการตรวจสอบความตรงด้วยการหาค่าความสัมพันธ์กับค่าที่คำนวณได้จากสูตรของบาร์คและโรเดนบุช (Bryk & Raudenbush) พบว่า มีค่า 0.98 สำหรับทั้งสองมหาวิทยาลัย ส่วนการตรวจสอบความตรงอีกวิธี ได้จากการสัมภาษณ์ความคิดเห็นของหัวหน้าภาคและตัวอาจารย์เอง เพื่อเปรียบเทียบค่าการเปลี่ยนแปลงที่ได้จากสูตรที่ผู้วิจัยพัฒนาขึ้นกับค่าที่ได้จากสูตรของบาร์คและโรเดนบุช พบว่า อาจารย์ร้อยละ 65 เห็นว่าค่าที่ได้จากสูตรที่ผู้วิจัยพัฒนาขึ้น ตรงกว่าค่าที่ได้จากสูตรของบาร์คและโรเดนบุช มีเพียงร้อยละ 10 ที่เห็นว่าค่าที่ได้จากสูตรบาร์คและโรเดนบุช ตรงกว่า ส่วนหัวหน้าภาคเห็นว่า ร้อยละ 62.5 ของค่าที่คำนวณได้จากสูตรที่ผู้วิจัยพัฒนาขึ้น ตรงกว่า มีเพียงร้อยละ 20 ที่เห็นว่าค่าที่ได้จากสูตรบาร์คและโรเดนบุช ตรงกว่า ดังนั้น ค่าการเปลี่ยนแปลงที่คำนวณจากสูตรสำหรับเส้นโค้งที่ผู้วิจัยพัฒนาขึ้น มีความตรง

ภาควิชา วิจัยการศึกษา
สาขาวิชา การวัดและประเมินผลการศึกษา

ลายมือชื่อนิสิต .....
ลายมือชื่ออาจารย์ที่ปรึกษา.....
ลายมือชื่ออาจารย์ที่ปรึกษาร่วมคนที่ 1 .....
ลายมือชื่ออาจารย์ที่ปรึกษาร่วมคนที่ 2 .....

4184972027: MAJOR EDUCATIONAL MEASUREMENT AND EVALUATION
KEY WORD: PERFORMANCE SCORES / UNIVERSITY INSTRUCTORS / CHANGE
JITLEKHA TEERAJARMORN: DEVELOPMENT OF METHODOLOGIES TO EVALUATE THE PERFORMANCE SCORES OF UNIVERSITY INSTRUCTORS: A CASE STUDY IN TWO UNIVERSITIES. THESIS ADVISOR: PROF.UTUMPORN JAMORNMANN, Ph.D. THESIS CO-ADVISOR: PROF.PETER ROWLINSON, Ph.D. AND ASSOC.PROF.ARUNEE ONSAWAD, Ph.D.

The objectives of the research were to develop a methodology for measuring university instructors' performance scores and to develop formulae for calculating performance change scores over a 3 year period.

Following a literature review and interviews with senior academic staff from Thammasat University and Saint John's University, 35 subtasks, grouped under six principal tasks, were proposed for the scoring of university instructors' performance. The subtasks were endorsed as being appropriate by further samples of senior and experienced staff from both universities. Additional samples of senior staff were consulted to determine the weight each task and subtask should contribute to instructor performance scores at each university. Many of the tasks and subtasks had similar weightings at both universities and the differences that did exist, whilst large enough to justify the use of separate sets of weighting factors, tended to follow the pattern that was expected given the two universities' different missions and backgrounds.

Data were collected for three academic years (1998, 1999 and 2000) for a total of 290 instructors from both universities. Data for 14 of the subtasks was obtained from existing university documents; a self-report questionnaire was developed for instructors to give data for the remaining 21 subtasks. Cronbach's alpha was around 0.8 each year for both universities, indicating the scores had acceptable internal reliability. When subtasks were ranked by their correlation with total score, there was a high degree of similarity in the rankings each year within each university, implying considerable stability. There was much less similarity in the ranking between the universities, implying that there were differences between the universities in the relationships between the subtasks and the total, even though the reliability of the scores was acceptable at both universities. Samples of Heads of Department from both universities were asked to state whether the scores for each of the instructors within their departments were valid. For both universities, 90% of the instructors considered had scores declared valid. The methodology therefore produced performance scores that were reliable and valid within each university.

Two formulae were developed for assigning scores to instructors based on the size and pattern of the change in their performance scores over the 3 year period. One formula was developed for linear change, the other for non-linear change. Since almost all the instructors (96%) for whom data were collected exhibited non-linear change, only the non-linear formula could be evaluated. Performance change scores calculated using the researcher's formula were validated by calculating the correlation between them and change scores obtained from the Hierarchical Linear Model (application of quadratic growth model) of Bryk & Raudenbush. The correlation was 0.98 for both universities. Further validation was obtained by asking heads of department and instructors to compare the change scores obtained from the researcher's formula with those obtained from Bryk & Raudenbush's method. 65% of the instructors considered their score from the researcher's formula to be better than that from Bryk & Raudenbush's formula, compared with 10% who thought Bryk and Raudenbush's better. Heads of department considered 62.5% of the researcher's scores to be better, and only 20% of Bryk and Raudenbush's to be better. There was therefore evidence that change scores calculated using the researcher's non-linear formula were valid.

Department Educational Research	Student 's signature ……
Field of study Educational Measurement and Evaluation	Advisor 's signature ……
Academic year 2002	Co-advisor 1 's signature ……
	Co-advisor 2 's signature ……

Acknowledgement

The dissertation could not have been completed successfully without the invaluable help of my advisor, Professor Dr.Utumporn Jamornmann. I would like to extend my deepest and sincerest gratitude for her guidance, encouragement, kindness and support throughout my research.

I would also like to express my gratitude to Professor Dr.Peter Rowlinson and Associate Professor Dr. Arunee Onsawad for their inspiration, helpful suggestions and support. My deepest thanks to my dissertation committee whose tireless efforts kept me on the right track.

Particular thanks are extended to Thailand Research Fund (TRF) for the scholarship that supported my study and research.

I am deeply grateful to all of instructors in my department of educational research for established knowledge. In addition, I owe much to senior staff, instructors and officers in Thammasat University and Saint John's University especially Dr. Suvichakorn Chinnapha for their useful cooperation in my research.

My personal thanks to Dr.Pensi Teeravarapaug, Mr. Nicholas J. Railes, Ms. Anchalee Leuang-on, Ms. Taneenart Na-Soontorn, Ms. Wanwimol Chumchusart and my friends in the Research and Evaluation Division of University of Cambridge Local Examination Syndicate and also all of my friends for their help and willpower.

Finally, I would like to thank my parents, Ms. Siri Teerajarmorn and Mr. Soontorn Teerajarmorn and my brother, Mr. Thammathiti Teerajarmorn for their love and willpower. To all of you, I offer my heartfelt thanks.

Jitlekha Teerajarmorn

CONTENTS

Page

CONTENTS (Continued)

Page

LIST OF TABLES

Page

LIST OF TABLES (Continued)

Chapter 1
Introduction

Research Problem

Quality is an essential issue dominating the debate in many countries about the outputs of higher educational institutions and whether societies are getting real value for their investment in higher education. Thailand is no exception (Ministry of University Affairs, 1998). The quality of educational development and academic excellence in Thailand was emphasized in the Eighth National Economic and Social Development Plan (1997-2001) and in the National Education Act of 1999, section 49-51, which stated that all institutions should develop their own internal quality assurance programmes and be ready for assessment by 'an Office for National Education Standards and Quality Assessment' (Office of the National Education Commission, 1999). An assumption behind this development was that the introduction of quality assurance programmes would raise quality.

Instructors (defined here as including Lecturers, Assistant Professors, Associate Professors and Professors) are a vital part of a university, so much so that there is a saying that the "professor is the university" (Boonprasert, 1999). Clearly, then, any substantial improvement in quality should be reflected in improvements in instructors' performances.

Many studies of Thai university instructors' performance pay scant attention to the quality and range of their activities. Most of the studies emphasize only quantitative, objective criteria, for example the number of hours worked, e.g. Umpuang (1985), Pittayanuwat et al. (1981), Boonying (1986) and Punsuwan (1994). Moreover, some research studied only teaching, such as the research by Buakam (1997) and Boondeekul (1998), even though instructors' tasks are much broader than teaching alone. In Thailand, there is a consensus that university instructors' duties fall under six main tasks: teaching, student advice, research and academic publications, acedemic service to community, preservation of arts and culture, and administration and academic self-improvement (Ministry of University Affairs, 1992; TU, 2000; SU, 1997; Tephatsadin N. Ayudtaya, 2000; Ubon Ratchathani University, 2001; Fujareon, 1988; Umpuang, 1985). Thus, it is necessary to have a valid and reliable way of scoring instructor's performance which covers all their major tasks and considers both quantity and quality.

For methodology to measure change, the International Encyclopedia of Educational Evaluation (1990) divided it into 2 main methodologies that are 'measuring change with 2 waves of data' and 'measuring change with multi-wave data' (Willet, 1990). Measuring change with 2 waves is used for calculating individual change i.e. difference change or gain score – difference of pre and post measurement, Webster and Bereiter's (1963) reliability-weight measure of change, Lord's (1956) regression-based estimated true change, Residual change scores – obtained by estimating residuals from the population regression of true final status on true initial status (Willet, 1990). It is true that pre/post (or two-wave) measurement is the most common design in the study of change, and two repeated observations do indeed constitute a longitudinal study. However, two observations are not adequate for studying the form of change. Two observations can only estimate the amount of change in straight line. Rogasa showed that if the rate of growth is not constant, but depends on time, the amount of change will depend crucially on the times of measurement, and observations of individuals at a difference set of two time points may give contradictory results (Gottman, 1995).

In the last decade, researchers have moved beyond the limitation of the pre/post or two-wave measurement. They have preferred and widely used the methodologies to measure change with multi-wave data. The methodologies to measure change for individuals could be classified into 4 groups: Structural Equation Model, Adopting Dynamic Theory, Applying Mathematics Principles, and Hierarchical Linear Model.

Structural equation models, for example Latent Curve Analysis by Meredith and Tisak (1990), which relies mainly on statistics hypothesis testing about the relation between observed variables and unobserved variables. Data analysis for Measuring change is on a condition that the number of respondents has to be at least 300 people (Bijlevled and et a. 1998), which may give problems for calculating change in a university that has a small group of instructors.

Measuring change adopting dynamic theory as it is in Multidimensional Latent Trait Model by Embertson (1991) to measure learning change. The results obtained from this model were the amount of individual change and quality of tests based on the basic assumption that examiners used their cumulative ability. It is also in Gultman Simplex Model by Collins and Cliff (1988). In this model, individual change and group change were found on the assumption that cumulative ability was used in the measurement situation, which is only in one direction and never decreased. If not, the measurement situation for research would not satisfy the forementional

assumption, and some variables e.g. instructor performance, may have many patterns of change.

Measuring change by applying mathematics principles such as Regression, Slope, Time series, or Polynomial Regression is simple method that is easy to calculate change score. But it cannot find the difference of change scores over a 3 year period in instructors who have different patterns of change, if they have the same scores in the first and third year of the measurement.

Hierarchical Linear Models were developed by Bryk & Raudenbush (1987). This methodology was considered to be a good methodology for measurement of change (Bijileveld *et al*, 1998; Khamlan, 1997; Ruachantuk, 1999) because it composed many statistics principles for calculating change (Bryk and Raudenbush, 1987). But it has a limitation in estimating parameters in the case of less than 100 respondents (Bijileveld and et al, 1998) and it is complex for general people because of its advanced methodology.

Therefore, It is necessary to have methodologies that are easy to calculate and interpreted and one can calculate change scores for individuals without impact of sample size.

Thammasat University (TU) and Saint John's University (SU) were selected for the case study. They introduced quality assurance in the academic year 1998 and participated in the pilot project for research and development "Indicators, Criteria, and Techniques for Internal and External Assessment in Thai University" run by Prof.Dr.Utumporn Jamornmann on June 2000 - March 2001. Both universities managed quality assurance at university and faculty level which included self-study, self-evaluation, internal audit, and internal assessment in 9 aspects: 1) philosophy/objectives/ implementation, 2) teaching and learning, 3) student development activities, 4) research, 5) academic service to community, 6) preservation of art and culture, 7) administration and management, 8) finance and budgeting, and 9) QA system and mechanisms. Any improvement in quality of either university should be reflected in improvements in instructors' performance especially in academic year 1998, 1999 and 2000 (a 3 year period) following the introduction of quality assurance in their university.

Thus, in this research a methodology to measure instructors' performance score and formulae for calculating performance change score over a 3 year period were developed. It is hoped that the findings of this research will have benefit for calculating university instructors' performance scores and change scores.

Objectives

The objectives of the research are as follows:

1. To develop a methodology to measure university instructors' performance scores

    1.1 To identify tasks and subtasks of university instructors

    1.2 To identify weighting of tasks and subtasks

    1.3 To calculate performance scores

2. To develop formulae for calculating performance change scores over a 3 year period

    2.1 To develop the formulae for calculating performance change scores

    2.2 To validate change scores using the formulae developed by researcher

Research Hypotheses

Hypothesis 1

Many studies of Thai university instructors' performance pay scant attention to the quality and range of their activities. Most of the studies emphasize only quantitative, objective criteria, for example the number of hours worked, e.g. Umpuang (1985), Pittayanuwat et al. (1981), Boonying (1986) and Punsuwan (1994). Moreover, some research studied only teaching, such as the research by Buakam (1997) and Boondeekul (1998), even though instructors' tasks are much broader than teaching alone. In Thailand, there is a consensus that university instructors' duties fall under the main tasks of teaching, student advice, research and academic publications, acedemic service to community, preservation of arts and culture, and administration and academic self-improvement (Ministry of University Affairs, 1992; TU, 2000; SU, 1997; Tephatsadin N. Ayudtaya, 2000; Ubon Ratchathani University, 2001; Fujareon, 1988; Umpuang, 1985). Thus, a methodology that is suitable for measuring instructor's performance should cover all major tasks and consider both quantitative and qualitative aspects. Besides this, the weighting of each major task may be different because of different university missions and backgrounds.

Therefore, "A methodology for scoring university instructors' performance that covers all their major tasks and considers both quantity and quality by multiplying the scores of subtasks developed with the weighting obtained from their university's experts and summing the weighted scores, should be an appropriate methodology."

Hypothesis 2

Many studies of change pay attention to learning variable, especially achievement. Ruachantuk (1999) studied Environmental knowledge at 7 time points, the result showed there was a linear change in Environmental knowledge. Willamson, Appelbuam and Epanchin (1991) studied achievement scores at 8 waves, the results showed only 38 individuals (7.2%) have quadratic change in reading and 26 individuals (4.9%) have quadratic change in Mathematics. Wijitwanna (2000) studied Mathematical abilities at 5 waves, the results showed there was a linear trend in calculation but there was a quadratic trend in the mean for problem solving. Tangsakulruanglai (1998) studied longitudinal changes in Mathematics achievement and physical development at 5 time points, the results showed Mathematics achievement was a downward quadratic while the curves of the students' weight and height were upward quadratic. Research of Khamlan (1997) found that the quadratic growth model could explain more variance in change of English vocabulary achievement at 8 time points than linear growth model.

For studies of change in other variables, they related to perception or behavior variables. Research of Chan et al. (2000) found that social skills development at home was best described with a nonlinear trajectory. MacCallum et al. (1997) illustrated the multilevel linear model of change using data from a study of physiological response to marital conflict in older couples from 5 blood samples at 5 time points, the result showed there was a linear trend in norepinephrine, but there was no systematic trend in the means of the other two hormones. Silverstein and Long (1998) studied about patterns of change in grandparent's perceptions of affection and in-person contact revealed quadratic trends in both growth curves.

From the above, researches has shown that almost all various variables had a non-linear or quadratic change.

Therefore, the performance scores change over a 3 year period (during the academic year of 1998, 1999 and 2000), should be non-linear.

Scope of the study

This study focused on the development of methodologies to measure university instructors' performance scores and the development of formulae for calculating performance change scores over a 3 year period. Thammasat University (TU) and Saint John's University (SU) were selected for the case study.

The scopes of the study are as follows:

1. Data were collected from 4 groups of academic staff in TU and SU as follows:

Group 1: Task Identification Group was composed of TU and SU senior staff (Academic Vice President, Deputy Dean, and Directors of University's Centers and Officers) and instructors with more than 5 years' experience.

Group 2: Experts for subtask weighting were composed of 17 senior staff. They have responsibility for quality assurance and were well aware of the instructor 's tasks.

Group 3: Target Group composed of TU and SU 's instructors who fit criteria.

Group 4: Validation Group was composed of 2 subgroups, (1) 13 heads of department and (2) 40 instructors.

2. The variable of the study was university instructors' performance, which covers 6 main tasks of instructor (teaching, student advice, research and academic publications, academic service to community, preservation of arts and culture, and administration and academic self-improvement).

Data for the 3 academic years 1998 (June 1998 – May 1999), 1999 (June 1999 – May 2000) and 2000 (June 2000 – May 2001) were collected.

3. A performance score was calculated for each instructor by multiplying each subtask score by the relevant subtask weighting factor for the instructor's university, and summing the weighted scores.

4. The performance change scores, using the formula developed by the researcher, were validated by the score changes obtained from the formula of Bryk & Raudenbush (1987). Further validation was obtained by seeking the opinions of heads of departments and instructors by comparing the score changes of individual instructors obtained by the formula developed by the researcher with those of Bryk & Raudenbush.

Basic Assumption

An instructor could recall the quality of his/her work in each of the academic years 1998, 1999 and 2000.

Limitation of the study

Data were collected only from TU and SU instructors in academic years 1998, 1999 and 2000 (a 3 year period).

Definition of terms in this study

Instructors The TU and SU instructors (define here as including Lecturers, Assistant Professors, Associate Professors and Professors) who had worked in the faculty since 1998, had never taken more than three months' leave in that time (e.g. personal leave, sick leave, maternity leave, sabbatical, etc.) and had never been a senior administrator such as university president, vice president, dean, deputy dean or director. There were 238 instructors from Thammasat University and 52 from Saint John's University who fitted these criteria and returned questionnaires.

Instructor's Tasks and Subtasks Six principal tasks consisting of teaching, student advice, research and academic publications, academic service to community, preservation of arts and culture, and administration and academic self improvement and 35 subtasks were identified to indicate instructors' performance.

Tasks and Subtasks' Weighting The weighting of each task at each university was the average of task weighting identified by their senior staff. The weighting for each subtask was calculated by multiplying the subtask average by the relevant task weighting and dividing by 100.

Instructor's Performance Score A performance score obtained by multiplying each subtask score by the relevant subtask weighting factor for the instructor's university, and summing the weighted scores.

Performance Change Score The university instructor's performance change score calculated by the formula developed by research following the consideration of pattern of change and reducing the floor and ceiling effect out off the observed change score.

Validity of Change Scores The change score calculated by formula developed by the researcher were validated by comparison with the score changes obtained from the formula of Bryk & Raudenbush (1987). Further validation was obtained by seeking the opinions of heads of department and instructors by comparing the score changes of individual instructors obtained by the formula developed by the researcher with those of Bryk & Raudenbush (1987).

Expected Outcomes

Expected outcomes of the study are as follows:

1. An appropriate methodology for measuring university instructors' performance score which can be used in other higher education institutions

2. Formula for measuring change scores which can be used for calculating change scores over a 3 year periods.

Chapter 2

Literature review

This chapter comprises 5 sections: (1) quality assurance in higher education, (2) tasks and subtasks of the instructors, (3) the development of indicators and weightings, (4) test of linearity and non-linearity and (5) measuring change.

2.1 Quality Assurance in Higher Education

2.1.1 Introduction

Higher education institutions are currently much involved with quality assurance. Quality assurance is about assuring interested parties that procedures are in operation to ensure quality (Cryer, 1993).

Much has been written on the meaning of 'quality' in Higher Education, and many definitions suggested, but the most commonly accepted is 'fitness for purpose'. This allows institutions to define their purpose in their mission and objectives, so quality is demonstrated by achieving these (Woodhouse, 1996).

In different countries, there are different approaches in quality assurance especially external quality assurance. These approaches can be grouped into 3 general categories: Accreditation (general accreditation/ specialized or professional accreditation), Quality audit and Quality assessment (Woodhouse, 1996).

Accreditation

This term has different meaning in various parts of the world. In the North American sense it can apply either to institutions or to programmes (subject or professional areas).

Accreditation assures the educational community, the general public, and other agencies or organizations that an institution or programme has clearly defined and educationally appropriate objectives, maintains conditions under which their achievement can reasonably be expected, is in fact accomplishing them substantially, and can be expected to continue to do so (Chermay, 1990).

It is noteworthy that in this definition of accreditation there is no requirement to judge whether the objectives (mission, aims) of an institution or programme are to meet any specified, or threshold standard. In many other countries, accreditation would imply that at least a threshold standard was intended and being achieved. For example, in the United Kingdom professional bodies accredit courses of study (programmes) meaning that graduates will be granted professional recognition (Frazer, 1991)

Accreditation encompasses both the objectives and the implementation of objectives: for example, it determines whether the objectives are appropriate for the institutional or degree level, and whether the resources are available to produce the desired outcomes. Accreditation is criterion-referenced; that is, it compares observed performance against preset standards usually determined by the accrediting agency. It generally involves a combination of performance indicators, self-study, and peer review. Accreditation may be performed at the institutional or program level, with program-level accreditation being most common in professional fields like accounting, business, law and engineering or for institutions offering degrees below the bachelor's level. The cycles are typically in the range of 10 years, unless serious problems are uncovered; such problems will lead to shorter cycle times or probationary status. The outcome, whether the institution meets threshold quality standards, is always published, such publication in necessary for accreditation to perform its certification function. However, details may be withheld to avoid adversarial relationships and, thus, to protect data acquisition and enhance accreditation's improvement agenda. (Dill *et al.*, 1996)

Quality Audit

The concept of quality audit has been developed in the United Kingdom, where in 1990 the Committee of Vice Chancellors and Principals established a small Academic Audit Unit using experienced academics on temporary secondment form universities. Quality Audit is neither concerned with a university's mission or objectives (inputs) nor with how successfully these objectives have been attained (outputs), but solely with the processes by which the university checks on the relations between its inputs and outputs. Sometimes quality audit is confused with accreditation (Frazer, 1991). Quality audit is like general accreditation, but without judging the suitability of the institution's objectives, which are taken as the starting point of audit (Woodhouse, 1996).

Audit is an externally driven peer review of internal quality assurance, assessment, and improvement systems. It focuses on the processes that are believed to produce quality and the methods by which academics assure themselves that quality has been attained. Audits of educational quality generally take place at the institutional level and focus on the formalities of quality assurance (on policy statements, rules and procedures, guidance notes, and minutes of meetings). Audits do not address academic standards, or determine the quality of teaching and learning outcomes, but evaluate how an institution satisfies itself that its chosen

standards are being achieved. The cycle times can be significantly shorter than assessment because audits are inherently simpler. (Dill *et al.*, 1996)

Quality Assessment

The process of assessment by peers of actual provision in particular subjects, which are often used in external quality assurance, is by scrutiny of institutional documentation and student work, direct observations, interviews, and by reference to performance indicators such as completion rates. Quality assessment is concerned with the outcomes of the process investigated by the internal quality audits, and normally requires that quality be measured or graded. While in principle quality assessment could, like quality audit, be at the level of the entire university, in England the Funding Councils are using it to assess specifically the quality of individual academic subjects or subject areas (Cryer, 1993).

The assessment process evaluates the quality of specific activities such as educational or research quality within academic units. Assessment goes beyond accreditation to make graded judgments about academic quality levels rather than binary judgments relative to threshold standards. Assessments generally are directed at the subject or program level, evaluating their delivered performance. Assessment uses a combination of performance indicators, self-study, and external peer review and defines quality relative to an institution's mission, not according to some universal standard of academic excellence to which only elite institutions can aspire. The cycle times tend to be in the range of 5 to 10 years. (Dill *et al.*, 1996)

Why we need quality assurance?

We need Quality Assurance in higher education because of 'accountability'. Accountability to at least three difference groups is depicted in the triangle in Figure 2.1 and the points discussed by Frazer are outlined below: (Frazer, 1991)

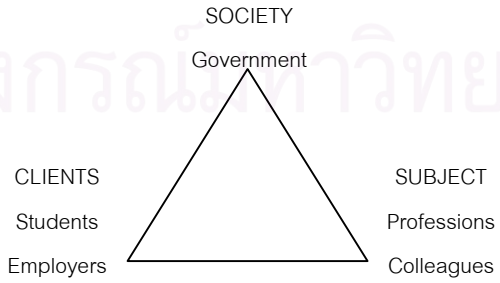


Figure 2.1. Accountability in Higher Education (Frazer, 1991)

Accountability to Society

In many countries there is a popular demand, and an economic necessity, for more higher education with consequential ever increasing costs, and in most countries society pays for much of this through taxes. Government acts for society in distributing funds to higher education either directly to the universities or indirectly through student grants or subsidized loans. Governments have a responsibility to society to ensure that what they 'buy' from higher education is acceptable and provides value for money. However, accountability to society is not only a matter of return on investment. Universities exist to safeguard and transmit a cultural heritage. Society needs assurance that universities are not failing in this duty.

Accountability to Clients

The clients of higher education are the students and the employers of graduates. They desire to have the best possible education available and then to receive certification that particular levels of knowledge and professional competence have been achieved.

Accountability to Subject

The third corner of the triangle is the subject. The knowledge, skills and attitudes that comprise each subject must not be distorted, suppressed or misused. Teachers are accountable to their professional colleagues that the integrity of their discipline is upheld and that students develop positive attitudes towards the subject and its use in society.

How can Quality Be Assured?

Real and enduring quality can only come by actions of the universities themselves. The basis for these actions must be 'self-evaluation'. Self-evaluation is not easy, but without three aids it is impossible. (Frazer, 1991) The first aid is external assistance. External agencies can provide external help for self-evaluation through the important ingredient of peer review. (Acherman, 1990) Peer should not only come from higher education; those actively engaged in industry, commerce and the professions must also be involved. The second aid is training (staff development) for the task of self-evaluation. Third, there is a need for national and international evaluation such as qualitative and quantitative performance indicators as well as descriptions of best practice and innovation in teaching, learning and assessment both general and subject-specific.

Quality assurance is the responsibility of the institutions themselves; real and enduring quality can only come from actions by the universities as a result of self-evaluation and peer review (Frazer, 1991).

2.1.2 Examples of Quality Assurance in Many Countries

2.1.2.1 The United Kingdom

Higher education in the United Kingdom has developed a quality assurance system, with teaching and research performances assessed separately.

1) Quality assurance in Teaching

Evaluation of quality assurance processes used in teaching began in 1990 with audit carried out by a body of the higher education sector, the Quality Assurance Group of the Higher Education Quality Council (HEQC). The former polytechnics and colleges were included in the audit after 1992 when they were given university status with the abolition of the binary system. In that year, to ensure that the quality processes in place lead to improvements in the standards of teaching and learning, the government introduced subject assessments, where the standard achieved in the teaching of individual subjects is evaluated. The Quality Assessment Committees of the Higher Education Funding Councils for England, Scotland and Wales conducted it. In 1997, the Quality Assurance Agency for Higher Education took over the two forms of evaluation and from 2001 it merged them into an integrated exercise. (Lim, 2001)

Quality Assurance Agency for Higher Education (QAA) is an independent body funded by subscriptions from universities and colleges of higher education, and through contracts with the main higher education funding bodies. Its mission is to promote public confidence that quality of provision and standards of awards in higher education are being safeguarded and enhanced. The current core business of QAA is to introduce an integrated quality assurance service by bringing together the academic quality audit of the HEQC and teaching quality assessment in each subject area by the Higher Education Funding Council for England, Scotland and Wales (which assessment in institutions in England, Scotland, Northern Ireland and Wales) (QAA, 2002).

QAA reviews the performance of higher education institutions at subject and institutional level *(the handbook for academic review are in www.qaa.ac.uk)* and publishes the finding of the reviews to provide public information and promote public understanding about higher education. It advises Government on degree awarding powers and university titles and manages the scheme for recognition of Access to Higher courses. Moreover, It works with the higher education sector to enhance quality and standards (QAA, 2002).

Academic Review

Academic review of QAA is the new, integrated method of review that focuses on the establishment, maintenance and enhancement of quality and academic standards. It has been used in Scotland since October 2000. Since January 2002, it has been in use across the whole of the UK. It operates over a six-year cycle, with each institution and all subjects being reviewed once in each cycle. (QAA, 2002)

The academic review process addresses three interdependent areas: (QAA, 2002)

- **reporting on academic standards** is concerned with the appropriateness of the intended learning outcomes (in relation to relevant subject benchmark statements, the qualifications framework and the overall aims of the provision); effectiveness of curriculum design and assessment arrangements (in relation to the intended learning outcomes); and the actual achievement of students;
- **reporting on the quality of learning opportunities** in a subject is concerned with the effectiveness of teaching, learning resources and academic support in promoting student learning and achievement;
- **reporting on institutional management of standards and quality** is concerned with the robustness and security of processes and procedures relating to the institution's responsibility as a body able to grant degrees and other awards that have a national and international standing. This involves, in particular, arrangements for dealing with approval and review of programmes, the management of academic credit and qualification arrangements, and the management of assessment procedures.

Making judgements (QAA, 2002)

- At subject level

For each subject area in an institution, a judgement is made about academic standards. Reviewers consider:

- whether there are clear learning outcomes that have been set appropriately in relation to the qualifications framework and any relevant subject benchmark statements;
- whether the curriculum is designed to enable the intended outcomes to be achieved;
- whether assessment is effective in measuring achievement of the outcomes;
- whether student achievement matches the intended outcomes and the level of the qualification.

In the light of this, reviewers will state whether they have:

- **confidence** in standards (a judgement made if reviewers are satisfied with current standards and with the prospect of those standards being maintained into the future); or
- **limited confidence** in standards (a judgement made if standards are being achieved but reviewers have doubts about the ability of the institution to maintain them into the future); or
- **no confidence** in standards (a judgement made if reviewers feel that arrangements are inadequate to enable standards to be achieved or demonstrated).

For each subject area reviewed in an institution, judgements about the **quality of learning opportunities** offered to students are made against the broad aims of the provision and the intended learning outcomes of the programmes.

Reviewers look at:

- effectiveness of teaching and learning - in relation to curriculum content and programme aims;
- student progression - recruitment, academic support, progression within the programme;
- learning resources - the adequacy and effectiveness of use of the library, equipment, accommodation and staff.

Each of these three categories is judged as either:

- **commendable** - provision contributes substantially to the achievement of the intended outcomes, with most elements demonstrating good practice; or
- **approved** - provision enables the intended outcomes to be achieved, but improvement is needed to overcome weaknesses; or
- **failing** - provision makes a less than adequate contribution to the achievement of the intended outcomes; significant improvement is required urgently.

Within the 'commendable' category, reviewers will identify any specific exemplary features that represent sector-leading best practice.

If reviewers have no confidence in the standards achieved or if they find that any aspect of quality of learning opportunities is failing, then the provision will normally be subject to a further formal review within one year.

- At institutional level

Institutional review addresses the ultimate responsibility for the management of quality and standards that rests with the institution as a whole. It is concerned particularly with the way an institution exercises its powers as a body able to grant

degrees and other awards. For example, it looks at institutional procedures for approval, monitoring and review of academic programmes; procedures for acting on the findings of external examiners, subject reviews, and other external scrutinizes; overall management of assessment processes; overall management of credit systems; and management of collaborative arrangements with other institutions.

It draws on the evidence of subject level reviews, and uses points of reference provided by sections of the *Code of practice. (*See *www.qaa.ac.uk)*

Reports list action points which are categorised as:

- **essential** - matters that are currently putting academic standards and/or quality at risk, and which require urgent corrective action;
- **advisable** - matters that have the potential to put academic standards and/or quality at risk, and which require either preventive, or less urgent, corrective action;
- **desirable** - matters that have the potential to enhance quality and/or further secure academic standards.

Judgements are made on the degree of confidence that may reasonably be placed in an institution's effectiveness in managing the academic standards of its awards and the quality of its programmes. A statement that confidence cannot be placed in institutional arrangements would result if there is a number of matters requiring 'essential' action, the combined effect of which is to render ineffective the quality assurance arrangements as a whole. A statement that limited confidence only could be placed in institutional arrangements might be made if there are a small number of 'essential' action points that could readily be implemented, or a large number of 'advisable' points. In all other cases there would be a statement of 'overall confidence' in the institutional arrangements.

The quality audit and the subject assessment are generally seen to have benefited teaching and learning. Together, they have succeeded in making universities more aware of the importance of having quality processes in place an in ensuring that the processes are implemented. (Lim, 2001)

2) Quality assurance in Research

The evaluation of research performance began in 1986, with the introduction of the Research Selectivity Exercise, which was repeated in 1989. It was replaced by the Research Assessment Exercise in 1992, which was repeated in 1996 and 2001. The assessment takes the form of assigning a rating on an institution’s research performance in designated subjects, with the rating scored and the number of

designated research-active staff used as the basis for allocating the main research funds for the higher education sector. (Lim, 2001)

The impact varies from institution to institution, though overall it is seen to have been positive. It has led to the development of a research strategy and culture and improved research management process. The need to designate research-active staff and protect the interests of the institution in the transfer market for high-flying researchers has also improved human resource management. (Lim, 2001)

Quality Assurance in the United Kingdom and their internal, external and funding impact on the institutions was summarized by Lim (2001) as showed in Table 2.1.

Table 2.1 Major Processes for Quality Assurance: United Kingdom (Lim, 2001)

Issue	Quality Audit	Teaching Quality Assessment	Research Assessment Exercise
Responsible agency	Higher Education Quality Council (HEQC) (institutions); moving to Quality Assurance Agency for Higher Education (QAAHE) (government and institutions)	Funding Council (government); moving to Quality Assurance Agency for Higher Education (QAAHE) (government and institutions)	Funding Councils (government)
Purpose	To support institutions' self-regulation by auditing the procedures by which they assure themselves of the quality of their academic provision	To ensure provision in of sufficient quality to justify public support to improve quality, and to "inform" funding and reward excellence	Highly selective distribution of funds in support of high quality research
Type of regulation (self, external, mixed)	Mixed	External	External
Scope (institution, department, program)	Institution	Subject area	Subject area
Activity assessed	Internal quality control mechanisms for teaching and learning	Teaching and learning	Research
Criteria (framework)	Nine broad aspects of institutions' quality control mechanisms	Six core aspects of subject provision (1995 onwards)	Research environment and plans
Standards	Mission-dependent	Mission-dependent	Adjusted national and international standards in each subject area
Evaluators	Predominantly peer review, with external assessors	Predominantly peer review, with external assessors drawn from the private sector and the professions	Predominantly peer review, with external assessors drawn from the private sector and the professions
Self-Study	Yes (self-criticism encouraged)	Yes (self criticism encouraged)	Yes (strengths highlighted; weaknesses downplayed)
Site visit	Yes	Yes	No
Indicators used	Predominantly textual material	Student entry profile, expenditure per student, progression and completion rates, qualifications attained, subsequent destinations	Peer-reviewed publications, research grant income, numbers of research assistants and students
Type of rating	Detailed written report, highlighting strengths and weaknesses	Each of six core aspects rated on a four-point scale (1995 onwards)	Seven categories, dependent on judgments concerning national and international standing
Dissemination	Funding Councils, institutions, potential consumers, press	Funding Councils, Web, institutions, potential consumers, press	Funding Councils, Web, institutions, potential consumers, press

Table 2.1 (Continue)

Issue	Quality Audit	Teaching Quality Assessment	Research Assessment Exercise
Financial impact	None	Funding withdrawn for persistent unsatisfactory provision; no reward so far (1997-1998) for excellent provision	Profound: core funding focused on research excellence
Internal impact	Significant; increasing: a more structured approach to quality control mechanisms	Significant; increasing: a more structured approach to the assurance of high-quality teaching and learning	Profound: organizational structure and management; faculty recruitment; teaching perhaps neglected
External impact	Modest; increasing: dissemination of best practice and reports on findings published	Modest; increasing: dissemination of best practice and reports on findings published	Considerable: bandwagon effect as more research sponsors, faculty, and students are attracted to strong areas

2.1.2.2 The New Zealand

The New Zealand Qualification Authority (NZQA) was established under the Education Amendment Act in July 1990 with a brief to oversee quality assured qualifications and to co-ordinate national qualifications in New Zealand (Barker, 1993). The Authority is responsible for the quality assurance of all education and training outside universities (NZQA, 2002a).

The Authority delegates quality assurance roles (other than degrees) in polytechnics and colleges of education to the Association of Polytechnics in New Zealand (APNZ) and the Association of Colleges of Education in New Zealand (ACENZ). Quality assurance in university programmes is the responsibility of the New Zealand Vice Chancellors' Committee (NZVCC) (NZQA, 2002a).

Approvals, Accreditation and Audit (AAA) is the NZQA business unit that provides an external check of the quality of the courses and qualifications in wänanga, industry, private and government sectors, as well as all degrees outside universities. In addition, the Authority registers unit and achievement standards, National Certificates, National Diplomas and other national qualifications on the **National Qualifications Framework**. Framework qualifications are quality assured and nationally recognised. (NZQA, 2002a)

Generally, AAA aims to protect the interests of learners; to make sure qualifications are meaningful and credible; to make sure qualifications are obtained in safe environments using appropriate teaching and assessment systems; and to assure the learner that NZQA-approved courses are well taught and nationally recognised. (NZQA, 2002a)

AAA verifies the quality of a provider's education programmes by checking their quality systems and effectiveness. It operates on a cost-recovery basis– providers pay for quality assurance activities. (NZQA, 2002a)

The role of AAA are to register private and government training providers; to accredit educational institutions and other registered learning providers to offer approved courses; to accredit educational institutions and other registered learning providers to award credits for qualifications registered on the Framework; to provide ongoing recognition of Industry Training Organisations (ITOs) as National Standards setting body and accredit ITOs to develop and monitor industry training programmes, and to register workplace assessors; and to approve courses and qualifications that are not based on registered national standards, including degrees not awarded by universities. (NZQA, 2002a)

Registration, Approval and Accreditation

Registration, approval and accreditation are the three main outcomes of the quality assurance process. The meaning of term was defined by NZQA as bellows: (NZQA, 2002b)

Registration

Registration indicates that a provider is capable of providing high quality education and training in a sound and safe learning environment. The Authority registers PTEs-approximately 860 are registered throughout New Zealand. Schools, polytechnics, universities, colleges of education and wananga do not need to be registered because they are set-up by the government under legislation.

Once registered, a PTE can apply for approval of its courses and accreditation (as below) and also apply for various forms of government funding.

Course approval

Course approval provides the public with an assurance that courses that are advertised "approved by NZQA" have been checked for quality.

An approved course is a coherent programme that is based on clear and consistent aims, content, outcomes and assessment practices. Some providers offer a wide range of courses, some of which will be approved by the Authority, and some of which might not be approved.

Accreditation

Accreditation signals that a provider is capable of running an approved course or awarding Framework credits.

A provider must be registered (or established under legislation) and accredited to be able to award credit for unit standards on the Framework, and also to: (NZQA, 2002a)

- receive programme funding from Skill New Zealand
- receive EFTS tuition subsidies from the Ministry of Education

- enable enrolled students to receive student allowances from the Ministry of Education
- deliver approved courses of three months or longer to foreign students.

When a provider first applies for registration, approval and accreditation, the provider must supply NZQA with a copy of its intended quality management system. This quality management system must describe whether the provider plans to set up as a single site or a multi-site organization and whether it will deliver face-to-face or by distance. NZQA will grant accreditation (assuming all other requirements are met) to the provider as either a single site or a multi-site provider, or as a classroom or distance deliverer, as appropriate. If a provider adds sites at a later stage, or changes its mode of delivery, then the basis on which initial accreditation was granted changes. (NZQA, 2002a)

How does NZQA assure quality?

NZQA takes a partnership approach to quality - it relies on three main processes: (NZQA, 2002b)

1. Ongoing monitoring of provider activity
2. Self-evaluation by providers
3. Quality audit.

Ongoing monitoring of provider activity

AAA provides ongoing monitoring of provider quality through liaison with other parts of the Authority (including Tertiary Records and Tertiary Assessment and Moderation), Industry Training Organisations and Government agencies.

Each quality assurance officer deals with a group of providers based on either geographical or provider type.

Self-evaluation

Providers are responsible for the quality of their own education and training programmes. They are expected to regularly undertake self-evaluation or internal review of their organisation and report their findings to the Authority prior to the NZQA audit.

Self-evaluation enables an organisation to self-assess their effectiveness against good practice criteria and to determine where they are in terms of compliance with the Authority's registration and accreditation requirements.

Self-evaluation enables an organisation to identify areas where improvement is needed and to develop action plans for improvement.

AAA provides a guide and workbook for organisations to evaluate themselves against requirements prior to audit. Providers can, of course, use their own self-evaluation tools.

Quality audit

An audit verifies the performance of a provider as a whole, their management processes for achieving quality learning, and their success as an education organisation. Each audit also involves a visit to the establishment.

Put very simply, quality audit asks the following questions:

- What education and training was planned and why?
- What resources and processes were put in place to do this?
- What actually happened? Was the programme run as planned and were the aims met?
- What was changed and why?

The quality audit is important in safeguarding the interests of existing and prospective students.

The standard for registration and/or accreditation

The Qualifications Authority standard for registration and accreditation – known as QA Standard One is the tool for AAA to decide whether a provider has a sound and stable learning environment. This leads to decisions on registration and accreditation. (NZQA, 2002a)

The provider has measurable goals and objectives for education and training. The provider demonstrates by stating its education activities within a written statement of its goals and objectives, approved by the governing body and having performance indicators to measure the achievement of its goals and objectives. (NZQA, 2002a)

The provider put into practice quality management systems to achieve its goals and objectives. The quality management systems must cover 8 aspects: 1) governance and management, 2) personnel, 3) physical and learning resources, 4) learner information, entry and support, 5) development, delivery and review of programmes, 6) assessment and moderation, 7) notification and reporting on learner achievement and 8) research. (See QA Standard One in www.nzqa.govt.nz)

The provider is achieving its goal and objectives, and can provide assurance that it will continue to do so. The provider demonstrates by applying suitable performance indicators to measure and monitor the achievements of goals and objectives; using the results of its performance measurement to update its goals, objectives and performance indicators; regularly collecting feedback from learners,

clients, funders and other stakeholders to confirm its effectiveness as a training provider and to further improve its education and training; and using review and evaluation processes to confirm that processes are being applied consistently and are effective in achieving desired outcome or to identify areas for ongoing improvement. (NZQA, 2002a)

2.1.2.3 The United States

The Council for Higher Education Accreditation (CHEA) is the new voice of the nation's colleges and universities on all matters regarding accreditation – a uniquely American approach to assuring quality and public accountability in institutions and programs through voluntary, non-governmental self-regulation. It is a private, nonprofit national organization that coordinates accreditation activity in the United States. CHEA represents more than 3,000 colleges and universities and 60 national, regional, and specialized accreditors. (CHEA, 2000a)

CHEA recognition of accreditation organization has 3 purposes that are to advance academic quality, to demonstrate accountability and to encourage purposeful change and needed improvement (CHEA, 1998).

When seeking recognition, the accrediting organization must address five CHEA standards that correspond to CHEA purposes as follows: (CHEA, 1998) (please see detail of standards in www.chea.org)

- *Advance academic quality*. Accreditors are required to have a clear definition of quality and clear expectations that the institutions or programs they accredit have processes to determine whether quality standards are being met.

- *Demonstrate accountability*. Accreditors are required to have standards that call for institutions and programs to provide consistent, reliable information about academic quality and student achievement to foster continuing public confidence and investment.

- *Encourage purposeful change and needed improvement*. Accreditors are required to encourage planning for purposeful change and scrutiny for needed improvement through ongoing self-examination in institutions and programs.

- *Employ appropriate and fair procedures in decision-making*. Accreditors are required to maintain appropriate and fair organizational policies and procedures that include effective checks and balances.

- *Continually reassess accreditation practices.* Accreditors are required to undertake self-scrutiny of their accrediting activities.

United States accreditation in higher education is a collegial process of self-review and peer review for improvement of academic quality and public accountability of institutions and programs. This quality review process occurs on a periodic basis, usually every 3 to 10 years. Typically, it involves 3 major activities: (CHEA, 2000a)

- A self-evaluation by an institution or program using the standards or criteria of an accrediting organization.
- A peer review of an institution or program to gather evidence of quality.
- A decision or judgment by an accrediting organization to accredit, accredit with conditions, or not accredit an institution/program.

Moreover, the other accrediting organizations are regional associations i.e. Middle States Association of Colleges and Schools (www.msache.org), New England Association of Schools and Colleges (www.neasc.org), New England Association of Schools and Colleges (www.neasc.org), North Central Association of Colleges and Schools (www.ncacihe.org), Northwest Association of Schools and Colleges (www.cocnasc.org), Southern Association of Colleges and Schools (www.sacscoc.org), Western Association of Schools and Colleges (www.wascweb.org), and Western Association of Schools and Colleges (www.wascweb.org) and also specialized and professional accrediting organizations i.e. Accreditation Board for Engineering and Technology, Inc., Accrediting Commission for Acupuncture and Oriental Medicine etc. (CHEA, 2000b)

2.1.2.4 Japan

The Japanese University Accreditation Association (JUAA), Inc., is an independent organization of universities. The Association was organized on July 8, 1947, under the sponsorship of 46 national, public, and private universities. The Association was established to "improve the quality of universities in this country by self-directed efforts and the mutual support of its members." Since 1956, the national government promulgated the University Establishment Standards via a ministerial ordinance, the Association's University Standards have been administered solely as standards for accreditation by this organization. During the 45 years that have passed since the 1951-52 academic year, when the first accreditation was granted, the Association has effectively administered its accreditation system in examining the qualifications of its formal members. Since 1996, the Association has launched a new accreditation system, one that incorporates into the examination and appraisal processes self-study by universities. (JUAA, 2001)

The accreditation of a university by the Association can take one of two forms: accreditation and re-accreditation. The former is applicable to a university that is to be newly admitted as a formal member of the Association, and the latter is granted thereafter in 10-year intervals to formal members that are accredited universities, after review by the Association. To be eligible for accreditation by the Association, a university must have been in existence for at least four years. (JUAA, 2001)

University accreditation by the Association serves the functions that benefit for universities as follows: (JUAA, 2001)

First, a university so accredited will be socially and publicly warranted to fill the role required of a successful university and to endeavor to improve and reform itself through self-study, with the aim of applying its own missions and realizing its own objectives.

Second, such a university is entitled to receive useful advice from the Association regarding how to further develop its strong points and how to correct its problem areas, with the aim of applying its own missions and realizing its own objectives.

Third, because one condition for accreditation by the Association is that a university seeking accreditation or re-accreditation monitors and evaluates itself in terms of vitally important requirements specified by the Association, a university can use this requirement to good advantage by creating and implementing a comprehensive program for self-study that really is suited to its own characteristics and features.

As the foregoing indicates, accreditation by the Association serves to certify the qualities of a university and to provide it with the opportunity to take advantage of such certification to more effectively make efforts to improve and reform itself.

The process of accreditation or re-accreditation is as follow: (JUAA, 2001)

First, a university seeking accreditation or re-accreditation by the Association will submit to the Association, not later than the specified date, statements setting forth the results of its self-study, as the case may be, with regard to the topics specified by the Association.

The 11 key items for self-study within the scope of the university organization and it activities are 1) missions and objective of the university and its colleges, 2) education and research organizations, 3) student admission, 4) education curricula, 5) research activities, 6) educational organization, 7) facilities and equipment, 8) books and other learning resources, and libraries 9) consideration given to student

life, 10) governance and administration, and 11) organization and systems for self-study. (*See* www.juaa.or.jp/english)

The statements shall consist of a Basic Institutional Data Report and a Self-Study Report. The Basic Institutional Data Report shall contain data concerning those matters relating to the university's organization and activities that can be demonstrated quantitatively. The required data will be compiled in accordance with the items and format specified by the Association. The Self-Study Report shall set forth the results of self-study conducted by the university, as well as its plans for future improvement, based on the data contained in the Basic Institutional Data Report and other qualitative data that might not be contained in the Basic Institutional Data Report. Items relating to self-study are specified by the Association in advance and are selected by the Association from among those items concerning the university's organization and activities that are deemed vitally important.

Second, before the Association reaches a decision regarding overall accreditation or re-accreditation, as the case may be, it examines, through its organizational bodies responsible for accreditation or re-accreditation, the statements submitted by the university.

The accreditation and re-accreditation processes concentrate, on a case-by-case basis, on, among other factors, the following: first, whether, based on a particular university's Basic Institutional Data Report, the university fully satisfies such minimum requirements as would be met by an formal member of the Association; and, second, based on the university's Self-Study Report, an evaluation of the efforts that are being made by the university or faculty to improve and to reform itself, with the aim of realizing its own missions or objectives.

Third, the Accreditation Committee makes the final recommendation regarding the accreditation of each particular university; and, likewise, the re-accreditation Committee makes the final recommendation regarding the re-accreditation of each particular university. In addition, the Accreditation Committee or the re-accreditation Committee, as the case may be, formulate proposals for the improvement or reform of each university. After approval by the Association's Board of Councilors and Board of Trustees, those recommendations and proposals are promptly sent to the university concerned.

In order to prepare itself to implement the new accreditation system that goes into effect in the autumn of l996, the Association has adopted an across-the-board amendment to its University Standard. It has been specifically indicated that the new University Standard are to be guidelines for universities to maintain and improve reasonable standards, and that they are to reflect the standards by which the

Association shall accredit and reaccredit. (*See* www.juaa.or.jp/english) The University Standards are set out in general and rather abstract terms. They also place emphasis on the missions and objectives of each university. In addition, it must be implemented with due respect to the uniqueness and special features of each university and promote the further development of these qualities. Thus, the university furthers to improve itself while maintaining its own uniqueness and its special features. (JUAA, 2001)

2.1.2.5 Hong Kong

Questions regarding the efficiency, cost-effectiveness, and economy of higher education are emerging, especially as the country is in an economic downturn and massive unemployment. In response to public concern, the Executive Council entrusted the University Grants Committee of Hong Kong (UGC) to start a quality assurance exercise at all of the government-funded higher education institutions. The UGC has made important progress in assuring greater value for money and the cost-effectiveness of higher education institutions' activities. A variety of mechanisms, including institutional and academic reviews, sectoral reviews, formal and informal visits, and discussion at various levels was begun in 1993 by the UGC. (Mok, 2000)

The UGC in its mission statement pledges to uphold the academic freedom and institutional autonomy of the institutions while at the same time seeking to assure the quality and cost-effectiveness of their education provision, and being publicly accountable for the public money devoted to higher education (French, 1997). The UGC has given attention to study and measure quality of the output (i.e. the graduates) (UGC, 1996). It believes that quality of higher education would be enhanced and maintained by performance indicators or quality assurance mechanisms (Mok, 2000).

The UGC contacted two Research Exercise (RAE) in 1994 and 1996 and one Teaching and Learning Quality Process Audit (TLQPR) in 1997, with the aim of evaluating if universities had properly institutionalized self-monitoring and self-evaluation. Management Reviews (MRs), in which roles, missions, academic objectives, resource allocation, planning, and financial process mechanisms of individual higher education institutions, have been conducted for 2 years. (Ibid and et al, 2000)

Teaching and Learning Quality-Process Review (TLQPRs) were used to assure value for money and quality of teaching and learning in higher education institutions. They are an externally driven meta-analysis of internal quality assurance, assessment and improvement systems (Massy, 1997). The TLQPRs were first

conducted by the UGC in 1996 at the University of Hong Kong, Chinese University of Hong Kong, Hong Kong University of Science and Technology, Hong Kong Baptist University, and Lingnan College. Similar reviews were organized at CityU of Hong Kong, Hong Kong Polytechnic University, and the Hong Kong Institute of Education in January 1997. The goal of the reviews is to examine whether these higher education institutions have their mechanisms and systems that can assure quality of teaching and learning. Major attention are given to 5 dimensions: 1) Curriculum design, 2) Pedagogical design, 3) Implementation quality, 4) Outcomes assessment, and 5) Resource provision. Moreover, the processes and sub-processes are supposed to be institutionalized by the higher education institutions in Hong Kong to facilitate learning and promote quality teaching, the reviews posed questions concerning four cross-cutting “meta-areas” that pertain to the institutions’ quality assurance and improvement environment (quality-program framework, formal quality program activities, quality-program support, and values and incentives). (Ibid and et al, 2000)

CityU was the first higher education institution in Hong Kong to establish a quality assurance committee (QAC) in March 1993 (Mok, 2000). CityU initiated an internal audit to review whether proper process and systems were in place to prepare for UGC-run teaching assessment exercise and pressed demand for research and publication by promulgating a strategic plan to help it excel in that area (Mok, 2000). It implemented a ‘Performance Planning, Appraisal, and Development’ scheme in 1995. The quality of performance of staff is certainly an individual as well as an institutional responsibility and appraisal is a process for quality assurance as distinct from quality control. The evaluation exercise is very comprehensive assessing faculty performance on different fronts, such as teaching, research, administration and contribution to the community and profession. In its management, CityU proposed a ‘merit-based reward scheme’ to reward exceptionally outstanding performance and sectioned colleagues with poor performance or underperformance. These initiatives clearly demonstrate how a ‘management-orientation’ approach that relies on quality assurance and audit, and efficiency has been implemented in a university. (Mok, 2000)

Most of the higher education institutions in Hong Kong in the territory have already undertaken the reduction of benefits for academic staff. In the past few years, there has been a general increase in academic teachers’ workloads and an increase in the faculty-student ratio from 1:12 to 1:20. Universities also respond to the reductions in government funding by raising research money from trusts or industry and by securing sponsorship for particular activities. Furthermore, Universities have

created consultancy firms and ventured into business and commercial activities to earn additional income to support educational development. Thus, University governance has shifted from the traditional collegial approach to a management-oriented and market model. (Mok, 2000)

2.1.2.6 Thailand

The quality issue is an essential issue which currently dominates the higher education debate in many countries about the outputs of higher education institutions and whether societies are getting real value for their investment in higher education. Thailand is no exception. (MUA, 1998)

The transition of the overall implementation began to be seen after the promulgation of the National Education Act of 1999. It is stipulated under Section 47 on 'Educational Standards and Quality Assurance' which reads: (MUA, 2001)

"There shall be a system of educational quality assurance to ensure improvement of educational quality and standards at all levels. Such a system shall be comprised of both internal and external quality assurance."

The Act under Section 49 also specifies that: (MUA, 2001)

"An office for National Education Standards and Quality Assessment shall be established as a public organization, responsible for development of criteria and methods of external assessment, conducting evaluation of educational achievements in order to assess the quality of institutions.

All educational institutions shall receive external quality assessment at least once every five years since the last assessment and the results of the assessment shall be submitted to the relevant agencies and made available to the general public."

Quality in Thai higher education is defined as a quality mechanism of production of graduates so that the public can gain confidence on the quality of education. Quality assurance in Thai higher education refers to activities as a mechanism, planning, and systematic activities implemented according to 9 aspects of higher education criteria: (1) philosophy, commitments, objectives and implementation plan, (2) teaching and learning, (3) student development activities, (4) research, (5) academic service to community, (6) preservation of art and culture, (7) administration and management, (8) finance and budgeting, and (9) quality assurance system and mechanisms (See MUA, 1998). It is a way to enable the public to be assured that higher education institutions are doing a competent job in ensuring quality productivity. (MUA, 1998)

Quality Assurance Policy

The quality of educational development and its academic excellence has been emphasized in the Eighth National Economic and Social Development Plan (1997-2001). The Ministry of University Affairs has played a significant role to encourage higher education institutions in maintaining and improving their academic standards so that they can produce graduates who satisfy the need of employers. In order to achieve the objective, the quality assurance policy was proclaimed for institutions to implement for better productivity on July 8, 1996. (MUA, 1998) The policy has stipulated that all universities improve and enhance their efforts for quality of instruction and the academic learning environment. The Policy Announcement is as follows: (MUA, 2001)

1. The Ministry of University Affairs will develop the quality assurance system and mechanism as an instrument to maintain institutional academic standards. The main principle is for all the higher education institutions to have quality control system and to consistently improve the performances of all their functions. Such implementation must be based on academic freedom and autonomy as well as public accountability for internationally recognized standards and heightened competitiveness. A subcommittee of education standards at the higher education level will be established to monitor and administer education standards and accreditation.

2. The Ministry of University Affairs will encourage institutions to develop their own internal quality assurance system in order to be a tool to improve quality of their educational management. The emphasis is to create an internal quality control mechanism of all the aspects influencing educational quality. This allows flexibility for each institution to set up its own internal audit and assessment systems as seen appropriate.

3. The Ministry of University Affairs will formulate principles and directions for the start up of the actual procedures. Each institution is able to make adjustments and improvement to fit in with its own conditions as so desired.

4. In order for the institutions to gain recognition for its internal quality assurance process by agencies in the wider circle and to demonstrate the quality of educational provision, the MUA will provide mechanisms for quality audits and assessment at the institutional and faculty levels before granting accreditation subsequently.

5. The Ministry of University Affairs will support and encourage both public and private higher education institutions including academic or professional associations to participate in quality assurance activities.

6. The Ministry of University Affairs will facilitate institutions to widely disseminate their information and the results of institutional quality assurance activities for the public to acknowledge higher educational standards. Such information will also be helpful for students and parents alike to make decisions on selecting desirable institutions. Additionally, it serves as a useful information source for the Government to consider the allocation of budget and resources for institutions which will further stimulate continuous quality improvement.

Processes of Quality Assurance

Since the promulgation of the National Education Act of 1999, the definitions of quality-related terms stated in the Act namely external quality assurance and internal quality assessment, have been used as the base for the operation by all the agencies and institutions involved. (MUA, 2001)

Internal Quality Assurance process means assessment and monitoring of the educational quality and standards of the institutions from within. Such assessment and monitoring are carried out by personnel of the institutions concerned or by parent bodies with jurisdiction over these institutions.

In order to reassure that higher educational institutions under its supervision are capable of delivering quality educational products, the MUA has introduced IQA process consisting of quality control, quality audit and quality assessment.

- **Quality Control** is the installation of system(s) and mechanism(s) under each of the quality factors in order to monitor the institutional implementation that meets with quality indicators set.
- **Quality Audit** is the process of studying and analyzing whether institutions have system(s) and mechanism(s) to monitor their quality control and whether they have taken actual actions and obtained results from such operations.
- **Quality Assessment** is the process of analyzing and comparing the implementation results with quality indicators and assessment criteria.

Each university and faculty has to set its own system and mechanisms for quality control, corresponding with its philosophy, mission and quality factors specified by the MUA. (MUA, 2001)

An important quality control mechanism is an office or a committee responsible for QA process at the university or faculty level. The line of coordination is as shown. (MUA, 2001)

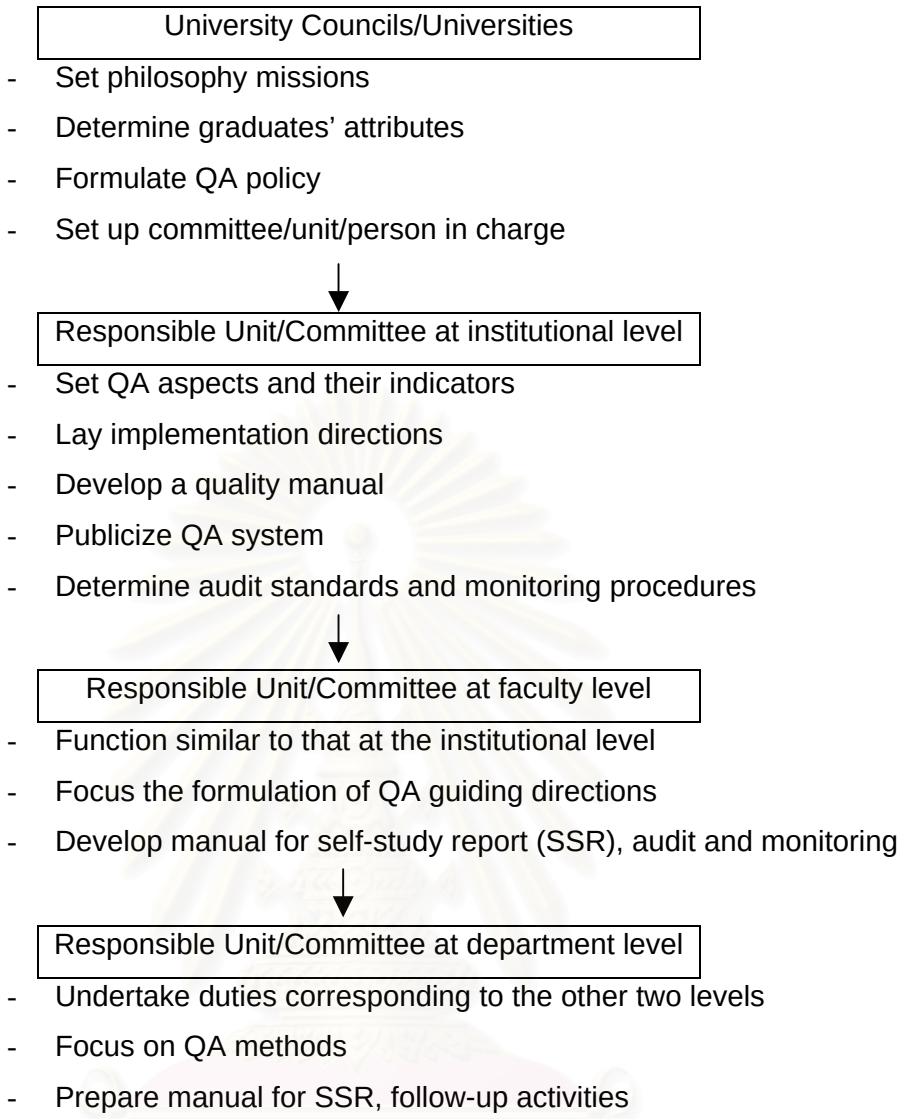


Chart 1: System and mechanisms for Quality control develop by institution (MUA, 2001)

Each university/institution is free to develop or adopt any quality assurance systems that are suitable to its own contexts. However, it should be accountable and transparent. (MUA, 2001)

External Quality Assurance means assessment and monitoring of the educational quality and standards of the institutions from outside. Such assessment and monitoring are to be carried out by the Office of the National Education Standards and Quality Assessment or by persons of external agencies certified by the Office. Such measures ensure the quality desired and further development of educational quality and standards of these institutions.

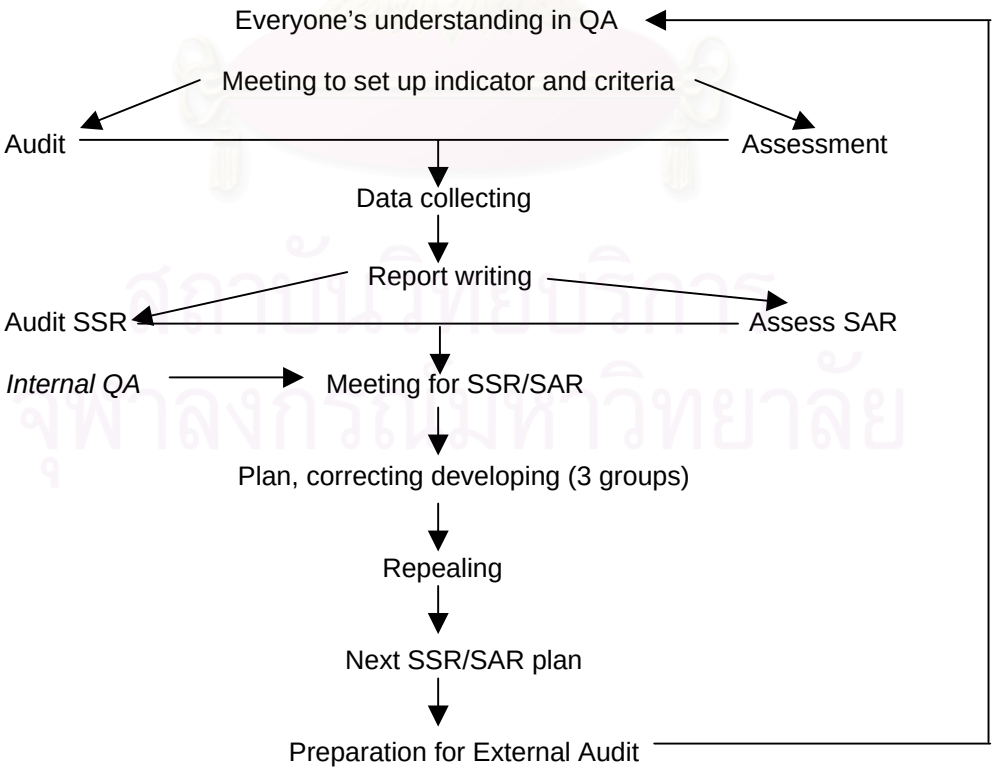
Table 2.2 The relationship between internal and quality assurance

Internal Quality Assurance		External Quality Assurance	
Organization	University/Institution The Ministry of University Affairs	Organization	The Office of the National Education Standards
Process	Quality Control Quality Audit Quality Assessment	Process	Quality Assessment

The National Education Act 1999 requires that each educational institution will receive external quality assessment at least once every five years and the assessment results will be submitted to the relevant agencies and be made available to the general public. The first round of external assessment of all education institutions will be complete by 20 August 2005. (ONEC, 2000)

Any type of Quality Assurance (QA) aims at achieving efficiency and effectiveness of all staff working with full capacity to produce quality products. Quality products of higher education institutes are (1) quality graduates who had success in meeting the curriculum objectives, (2) quality research from which research results, useful to both academics and the profession are disseminated, (3) quality academic service for society which is useful for both parties: givers and receivers, and (4) quality preservation of arts and culture which is impressed by organizer and participants (Jamornmann, 2001).

The circle of QA is presented below (Jamornmann, 2001).



Note SSR = Self Study Report, SAR = Self Assessment Report

2.1.3 Literature Related to Quality Assurance Concerning Instructor's or Academic Staff's Performance

Lim (1999) studied quality assurance in higher education in developing countries. Theoretical and empirical studies showed that output growth in developing countries depends importantly on the availability of workers with relevant and quality skills. They also show that technical progress becomes a more important determinant of output growth as economic growth proceeds. The experiences of universities in developed countries showed that quality assurance programmes improve the quality of their academic activities significantly only when academic staff are paid enough to live comfortably on their university salaries, have research skills, receive adequate support service, enjoy academic freedom, and are promoted for performing their duties well and not for having political or social connections, and do not work in environments where acrimonious personal relationships abound.

While many of the conditions required for the successful implementation of quality assurance programmes are not present in most universities in developing countries, their adoption will still be useful. They showed how a university's seemingly disparate activities are related to one another to serve a common cause and how the quality of these can best be improved by adopting an integrated approach. In the process, they provide more focus and direction to the work of the traditional academic committee system established to improve the quality of a university's work. However, the quality assurance programs must be modified to suit the conditions prevailing in developing countries. This requires that they should be kept simple in design, modest in the expectations created, and realistic in the resources required for implementation.

Wise (1996)'s study showed American schools remain geared to produce students for a country that, in many respects, no longer exists. Schools of education must provide school systems with prospective teachers who are able to help students perform in America's new society. As a result, teachers themselves must acquire new skills and new knowledge. In order to guarantee quality in teaching, the teaching profession must develop and embrace a system of quality assurance that is already employed by other professions. Accreditation, state licensing, and board certification can, together, create a coherent system for quality assurance for the teaching profession.

Dill (2000) studied academic audit in UK, Sweden, New Zealand and Hong Kong. The greatest lesson to be learned from the experience with academic audits in these countries is the accumulating evidence of its impact on systems of higher

education. In every case the implementation of academic audit has acted as a catalyst for change. Academic audit has helped initiate or bolster development of quality assurance systems within institutions, has placed attention to improving teaching and student learning on institutional agendas, has helped to clarify responsibility for improving teaching and student learning at the individual, academic unit, faculty and institutional level and has reinforced institutional leaders in their efforts to develop institution-wide 'quality cultures'. Moreover, they have facilitated discussion, cooperation and development within academic units with regard to means for improving teaching and student learning and have provided system-wide information on best practices and common problem areas, and also have offered visible confirmation to the public that attention is being paid to academic quality assurance. At a minimum, an academic quality assurance system that helps to build the capacity of academic institutions to maintain academic standards and improve the quality of teaching and student learning in the new competitive context appears to be needed and to be in the public interest.

Fourie & Alt (2000) studied the challenges to sustain and enhance the quality of teaching and learning facing academic staff in South African University. It has been argued that the best option for staff who find themselves in constraining conditions is to make a shift in emphasis form a resource-led culture to a problem-solving, improvement-led culture. In this regard, action research provides a theoretical framework for real improvement in teaching and could be seen as a good starting point to promote a culture of learning in the teaching and learning function of the university. However, it is a substantial task to move from informal and uneven departmental self-evaluation processes to institution-wide quality assurance systems. Furthermore, there are at least eight factors at work that interfere with academic staff's willingness and ability to respond to this quality agenda. However, the analysis has identified ways of advancing this quality agenda, but these come at a cost. More serious is the suggestion that if academic staff become occupied by buiding and conforming to formal quality assurance procedures, their attention may be diverted from teaching and research. In that sense, quality assurance that is not integrated into the core activities of academic staff, such as programme planning and development and professional growth and development, may harm quality.

Carroll (1997) 's research showed there are distinct challenges in pursuing quality advancement that result from the organizational structure and environmental context of universities. Massey University undertook a self-audit of quality during 1995 and 1996 with particular reference to academic and service departments. In

using this approach, two factors were discovered to have a major influence on pursuing quality advancement: the internal economic system and the concept of professionalism, the affiliation to a professional organization was interpreted as being synonymous with providing assurances of quality.

Hughes, Williams & Ryall (2000)'s research showed the findings of a wider research programme into the effectiveness of quality assurance implementation conducted at the University of Glamorgan. The study investigated the different approaches used towards the implementation of quality management systems in order to achieve registration to ISO9000. The poor quality staff indoctrination prior to implementing a change would appear to explain a substantial proportion of resistant behavior and that often too little attention has paid to staff indoctrination. The successful organizational change, which is planned must, be firmly based on a well-conceived and structured indoctrination strategy that aims to maximize staff awareness of the need for change and that the benefits of this will accrue to all. Management's role in change indoctrination is to convince their staff and, most importantly, their consultants of the benefits of close cooperation in change management. Consultants must be indoctrinated to adopt the correct approach. This presupposes active involvement of senior management in the implementation process and their full commitment to the process as part of a strategy for quality improvement. Teamwork must become a reality, not a catchword.

Brennan, Frederiks & Shah (1997) 's research considered the nature of the impact of quality assessment of the Higher Education Funding Council for England. The results showed quality assessment has led to the recognition of the need to have well developed and effective internal quality assurance procedures. The main rationales for making changes to institutional quality assurance procedures were to provide better preparation for external quality assessment, and also to lessen the burden of quality related activities on staff by reducing the time and resources spent on such activities. However, on the down side quality assessment has used up a lot of time and resource and caused a lot of stress. More positively, it has provided an impetus for institutions to give more attention to the quality of their teaching.

Newton (1999) 's research was an evaluation of the impact of external quality monitoring on a higher education college between 1993 and 1997. The particular focus of the study was on issues arising from the development and implementation of a quality assurance system at the study institution, and attempted to reconcile requirements for accountability and efforts to encourage quality improvement. The results showed that the improvement in quality for staff that the quality system

brought led to improvements in quality for staff, 73% of academic managers agreed but only 46% of all respondents agreed.

2.2 Tasks and Subtasks of University Instructors

From the history of higher education development in western countries, it appeared and was confirmed that the instructors (i.e. Lecturer, Assistant Professor, Associate Professor and Professor) are the core of university evolution. The instructor's roles, duties and performance reflect university's management. If one looks back to the 13th-14th century and the science revolution, one finds that there was an initiative for universities to be a place of acquisition of knowledge and transmission of knowledge. The community also agreed. Through out the 14th century, universities taught basic knowledge and science and had an academic atmosphere (Boonprasert, 1999).

Germany was the first country in which the universities paid attention research during the industrial revolution. Thus, the instructors in Germany were interested in research and did research, especially basic research but they abandoned teaching and community service. While Cambridge University and Oxford University, which were the main higher education institutions in England, emphasized on teaching and learning, and producing bachelor's students. Research was a proper activity for gifted persons (Ashby, 1964 referred in Boomprasert, 1999). In the United States, education for bachelor's degrees followed the English education system that aimed on produced the "whole-man", but education for graduate students followed Germany's educational system that emphasized advanced studies and research. Beside these, the US developed its modern educational system by integrating Bemjamin Frankin 's concept on education for humanity and Thomas Jefferson's which suggested a broad curriculum management. Thus, the main duty of the instructors in US was service to the community by managing teaching and research suitable for community needs (Boonprasert, 1999). In conclusion, there are 3 tasks for university instructors in western countries: teaching, research, and community service.

Higher education institutions in Thailand have important roles in training students to be good citizen of the community, to be a person who has traits and characteristics which the community wants; to be a resource person who is full of knowledge and to be a person who preserves arts and culture. Therefore, the main objectives of higher education institutions are comprised of 4 aspects: teaching,

research, academic service to community and preservation of arts and cultures (Ministry of University Affairs, 1998).

In Thailand, there is a consensus that university instructors' duties fall under six main tasks: teaching, student advice, research and academic publications, academic service to community, preservation of arts and culture, and administration and academic self-improvement (Ministry of University Affairs, 1992; TU, 2000; SU, 1997; Tephatsadin N. Ayudtaya, 2000; Ubon Ratchathani University, 2001; Fujareon, 1988; Umpuang, 1985).

1) Teaching

Teaching is the task that instructors usually spend most of their working time on (Suwanpat, 1990; Sapbamrung, 1994). Teaching is comprised of preparation, teaching, student assessment, and teaching evaluation (Jamornmann, 1998).

Teaching preparation is important for higher education instructors (Tephatsadin N. Ayudtaya, 2000). Instructors should have knowledge of their assigned subjects (Tephatsadin N. Ayudtaya, 2000; SU, 1997; Meesook, 2000) and should keep up to date with them (SU, 1997; Baer et al., 1996). They should work on 'lesson plans' (SU, 1997; Thaksin University, 1999; Bureau of Higher Education Standards, 1996; Baer et al., 1996) and also 'course outlines' (TU, 2000; SU, 1997; Tephatsadin N. Ayudtaya, 2000; Thaksin University, 1999; Jamornmann, 2000b) to give the students on the first day of teaching (Tephatsadin N. Ayudtaya, 2000; Thaksin University, 1999; Jamornmann, 2000b). The course outline should consist of the objectives and guidelines for study (Phethchuai, 1998; Sirichana, 1994). Any notes, etc. handed out by the instructors should be pitched at the appropriate education level and should be of a length appropriate to the learning period (Thaksin University, 1999; Sapbamrung, 1994). Instructional media used should be prepared in advance (Soliman & Soliman, 1997; Baer et al., 1996).

The 'number of teaching hours per week' is often used to indicate teaching workload (TU, 2000; SU, 1997; Thaksin University, 1999; Baer et al., 1996). The Thai Ministry of University Affairs has set the teaching workload for university instructors as follows: for teaching of graduate students only, the teaching hours per week should be not more than 6 hours and not more than 2 subjects per semester; for teaching of undergraduate students and graduate students, the total teaching hours for both levels should not more than 9 hours and not more than 3 subjects per semester; for teaching of undergraduate students, the teaching hours should be not more than 12 hours and not more than 3 subjects per semester. Also, 2-3 hours per

week of a practical subject is equal to 1 hour per week of a theoretical subject (Ministry of University Affairs, 1992).

Instructors should have teaching methods which are appropriate to the course's objectives, content, and students (TU, 2000; SU, 1997; Tephatsadin N. Ayudtaya, 2000; Phethchuai, 1998; Sirichana, 1994). The methods should be suitable given the backgrounds and experiences of the students and the emphasis should be on student learning (Phethchuai, 1998; Cave et al., 1997; Donaldson, 1994). Learning experiences should be provided through many different methods (Sirichana, 1994). Students should be introduced to relevant sources and reference books (Punyakanok and Tisyakorn, 1983), and given opportunities to present their studies (Phethchuai, 1998). Moreover, the instructors should provide activities and an atmosphere that stimulate students' learning (TU, 2000; SU, 1997; Tephatsadin N. Ayudtaya, 2000; Meesook, 2000; Bureau of Higher Education Standards, 1996; Sirichana, 1994). Instructors should be on time for both the beginning and end of teaching periods (SU, 2000; Wannakairoj, 1995).

Criteria for student assessment should be decided in advance (TU, 2000) and announced to students (TU, 2000; SU, 1997; Punyakanok and Tisyakorn, 1983; Donaldson, 1994; Sirichana, 1994). Instructors are responsible for developing tests and examinations (Baer et al., 1996), administering them at the right time (SU, 1997), and submitting students' grades to their faculty offices on time (TU, 2000). Instructors should use many different assessment methods (Tephasadion N. Ayayudthaya, 2000; Punyakanok and Tisyakorn, 1983).

Teaching evaluations consist of student evaluation (TU, 2000; SU, 1997; ; Phanphruk, 1995; Duangmanee, 1997; Donaldson, 1994; Cave et al., 1997), expert or employer evaluation (Cave et al., 1997; Donaldson, 1994), and peer or senior staff evaluation (Jamornmann, 2000a; Phethchuai, 1998). The student evaluation and self-evaluation are the basic information that is necessary to develop teaching quality (Bureau of Higher Education Standards, 1996). The instructors should use the evaluation results to improve teaching (TU, 2000; SU, 1997; Jamornmann et al., 2001) or analyze them for application (Donaldson, 1994).

2) Student Advice

Universities should be very supportive of giving students advice (such as academic advice, thesis advice) because it is a good mechanism for developing students (Tephasadion N. Ayayudthaya, 2000).

Instructors should be ready to give advice and be helpful to students (SU, 1997), and should give appropriate and adequate counselling or advice in academic,

career, and personal issues (Donaldson, 1994; Soliman & Soliman, 1997). Academic advice means explaining about choosing or withdrawing from courses, requirements of the curriculum, and details of scholarships, etc. (SU, 1997; Archava-amrong, 1994; Jamnongrak, 1989; Tephasadion N. Ayayudthaya, 2000; Baer et al., 1996), and warning when a student's grade is going down (SU, 1997; Archava-amrong, 1994; Jamnongrak, 1989). Moreover, the advisor should counsel and help students to solve their problems such as personal, social, family, and health problems (SU, 1997; Archava-amrong, 1994; Tephatsadin N. Ayudthaya, 2000; Baer et al., 1996). Instructors should give students information that is useful to them (Archava-amrong, 1994; Jamnongrak, 1989; Tephatsadin N. Ayudthaya, 2000), contacting the relevant office for information whenever he does not know the answers (Archava-amrong, 1994; Baer et al., 1996). They should set aside and inform students of specific times each week when they will be available for advice (Tephasadion N. Ayayudthaya, 2000; SU 1997; Archava-amrong, 1994; Takin University, 1999; Wannakairoj, 1995).

Relevant quantitative indicators for instructors' contribution to student advice include the 'number of advice hours per week' (Jamornmann et al, 2001; Umpuang, 1985) and 'number of advisees' (TU, 2000; Umpuang, 1985; Thaksin University, 1999; Cave et al., 1997). Thailand's Ministry of University Affairs set the ratio of students per instructor for undergraduate students on areas of specialization as follows: 1) Medical Science, Pharmacology, Fine Arts and Applied Arts, and Architecture – 8:1, 2) Engineering, Mathematics, Computer Science, Agriculture, Forestry and Fishery – 20:1, 3) Humanities, Philosophy and Religion, Social Science and Behavioral Science, Commerce and Business Administration, Mass Communication and Journalism, and Service Business – 25:1 and 4) Law – 50:1; the ratio for graduate students should be 5:1 (Ministry of University Affairs, 1992).

3) Research and Academic Publications

This is another major task of instructors and covers conducting research and writing textbooks, other academic books, academic articles, teaching supplement materials, and other innovations (Ministry of University Affairs, 1992).

Quantitative indicators for this task include the number of published research papers (Jamornmann et al., 2001; Duangmanee, 1997; Wannakairoj, 1995; Thaksin University, 1999: Duhs, 2000; Cave et al., 1997), the amount of research funds held (TU, 2000; Phanphruk, 1995; Duamgmanee, 1997; Cave et al., 1997) or the amount of research funds received from outside (Viriyawetkul, 1998; Chaiteeranuwatsiri, 1990; Soliman & Soliman, 1997), the number of awards (Phanphruk, 1995;

Viriyawetkul, 1998), the number of textbooks and other academic books written (TU, 2000; Duangmanee, 1997; Chaiteeranuwatsiri, 1990; Soliman & Soliman, 1997; Cave et al., 1997), the number of academic articles written (TU, 2000; Jamornmann et al., 2001; Ubon Ratchathani University, 2001) and published in academic journals (Duangmanee, 1997; Thaksin University, 1999; Chaiteeranuwatisiri, 1990) or presented in academic conferences (Chaiteeranutwatsiri, 1990; Soliman & Soliman, 1997) and the amount of citations (Cave et al., 1997).

Thailand's Ministry of University Affairs provides the following classifications for academic publications such as textbooks and other academic books: (1) good: contents are academic, correct, complete, modern, clear, and useful, (2) very good: the same criteria as good (level 1), moreover the publication presents knowledge or methodology that are useful and innovative and can be used for reference or practice, (3) excellent: the same criteria as very good, moreover it presents creative work, stimulates thinking and investigation, and is accepted at a national or international level (Ministry of University Affairs, 1992).

The Ministry also provides classifications for research: (1) good: appropriate methodology and analysis and a report showing academic progress that can be published, (2) very good: same criteria as good, moreover it presents significant new knowledge that can be used in academia or applied elsewhere, (3) excellent: same criteria as very good, but the work is creative, valuable, progresses knowledge and is accepted at a national or international level (Ministry of University Affairs, 1992).

4) Academic Service to Community

There are many ways an instructor can give academic service to the community (Chaiteeranuwatsiri, 1990), including 1) giving invited lectures (in or out of university), 2) being a visiting lecturer, 3) disseminating knowledge to the wider community by mass communication, 4) giving academic advice to the government or private organisations, and 5) arranging academic activities such as professional meetings, in-service training, seminars, and exhibitions (Sapbamrung, 1994; Suputsophon, 1992; Punsuwan, 1994; Ubon Ratchathani University, 2001).

A number of quantitative variables fit under this task, such as the number of invited lectures given (TU, 2000; SU, 1997; Jamornmann et al., 2001), number of hours per week devoted to academic service (Jamornmann et al., 2001; Ubon Ratchathani University, 2001; Supantat, 1984), and the number of times the instructor serves outside the university (Jamornmann et al., 2001).

Instructors' service to the community should be in their areas of specialisation (TU, 2000; Chaiteeranuwatsiri, 1990; Supantat, 1984) because knowledge should be

a base of academic service (Viriyawetkul, 1998; Fujareon, 1988). Furthermore, the services should fit the faculty plan or policy and be allowed from their faculty (Supantat, 1984), and also should satisfy the persons who received them (Wannakairoj, 1995; Chaiteeranuwatsiri, 1990).

5) Preservation of Arts and Culture

Instructors should establish morals, ethics, and culture in students (SU, 1997; Viriyawetkul, 1998; Fujareon, 1988; Suwanpat, 1990) and should be a good role models for their students (Tephatsadin N Ayudhiaya, 2000; Sapbamrung, 1994; Suwanpat, 1990). They should introduce or insert morals, ethics, and culture in teaching and learning methodology (SU, 1997; Archava-amrong, 1994), and should constantly participate in and support art and cultural activities (TU, 2000; SU, 1997; Suputsophon, 1992, Wannakairoj, 1995). They should occasionally organise projects that help preserve arts and culture (Thaksin University, 1999).

6) Administration and Academic Self-Improvement

This task covers everything else that the university requires instructors to do.

Variables relating to administration include attendance at committee meetings (Jamornmann et al., 2001; SU, 2000; Wannakairoj, 1995), the number of times attended at meetings (Jamornmann et al, 2001), the number of temporary and permanent committee memberships held (Jamornmann et al., 2001; Umpuang, 1985) and also the keeping of systematic evidence of work such as files, a database system, a reference system etc. (Fujareon, 1988).

Regarding academic improvement, an indicator of the effort put in is the number of times an instructor attends training courses, seminars or conferences relevant to his duties (SU, 1997; Duangmanee, 1997). Instructors should disseminate what they learn to other instructors in their department/faculty (Thaksin University, 1999) or implement the knowledge in their teaching or research (SU, 1997; Thaksin University, 1999).

Table 2.3 A list of subtasks for university instructors

Tasks and Subtasks	Subtask Type
Task 1: Teaching	
1.1 Has knowledge of assigned subjects	Quality
1.2 Keeps up to date with assigned subjects	Quantity
1.3 Course outlines produced	Quantity
1.4 Lesson plans produced	Quantity
1.5 Gives course outlines to the students in the first day of the course	Quantity
1.6 Course outline consist of the objectives and guidelines for study	Quantity
1.7 Any notes, etc. hand out were pitched at the appropriate education level and length of learning period	Quality
1.8 Instructional media prepared in advance	Quantity
1.9 Number of teaching hours per week	Quantity

Table 2.3 (Continue)

Tasks and Subtasks	Subtask Type
1.10 Teaching methods are appropriate to the course's objectives, content and students	Quality
1.11 Teaching methods are suitable given the backgrounds and experiences of the students and the emphasis is on student learning	Quality
1.12 Learning experiences were provided through many different methods	Quality
1.13 Students were introduced to relevant sources and reference books	Quantity
1.14 Students were given opportunities to present their studies	Quantity
1.15 Keeps up to date with assigned subjects regularly	Quality
1.16 Provides activities and environment that stimulate students' learning	Quality
1.17 Appropriate use of instructional media	Quality
1.18 On time for both the beginning and ending of teaching periods	Quantity
1.19 Criteria for student assessment were decided in advance	Quantity
1.20 Criteria for student assessment were announced to students	Quantity
1.21 Use many different assessment methods	Quantity
1.22 Examination administered on time	Quantity
1.23 Developing tests and examinations	Quantity
1.24 Grades reported to faculty on time	Quantity
1.25 Results of evaluation by students	Quantity
1.26 Self-evaluation in teaching	Quantity
1.27 Teaching evaluation by peer, senior staff and students	Quantity
1.28 Use of evaluation results	Quality
Task 2: Student Advice	
2.1 Number of advice hours per week	Quantity
2.2 Number of advisees	Quantity
2.3 Give counseling or advice in academic, career and personal issues	Quality
2.4 Specific times set aside for advising students	Quality
Task 3: Research and Academic Publications	
3.1 Number of published research papers	Quantity
3.2 the amount of research funds held	Quantity
3.3 the amount of research funds received from outside	Quantity
3.4 Number of awards	Quantity
3.5 the amount of citations	Quantity
3.6 Number of text written	Quantity
3.7 Number of academic books written	Quantity
3.8 Number of academic articles published in academic journals	Quantity
3.9 Number of academic articles presented in academic conference	Quantity
3.10 Classification for academic publications	Quality
Task 4: Academic Service to Community	
4.1 Number of invited lecturers (in or out university)	Quantity
4.2 Number of times serve outside the university	Quantity
4.3 Number of service hours per week	Quantity
4.4 Academic service to community on areas of specialization	Quality
4.5 Academic service to community fits into faculty plan or policy	Quality
4.6 Academic service to community was allowed from faculty	Quantity
4.7 Satisfaction of the persons who received service	Quality
Task 5: Preservation of Arts and Culture	
5.1 Establishes morals, ethics, and culture into students	Quality
5.2 Introduces or inserts morals, ethics and culture in teaching and learning methodology	Quality
5.3 Be good role models for students	Quality
5.4 Participant constantly in arts and culture preservation activities	Quantity
5.5 Participates and support art and cultural activities	Quantity
5.6 Organizes projects that help preserve arts and culture occasionally	Quality
Task 6: Administration and Academic Self-Improvement	
6.1 Number of times attended at meetings	Quantity
6.2 Attendance at department/faculty/ university meetings	Quantity
6.3 Number of permanent committee memberships	Quantity
6.4 Number of temporary committee memberships	Quantity
6.5 Keeps systematic documentary evidence of work undertaken	Quality
6.6 Number of times attended academic training/seminars/conferences	Quantity
6.7 Training/seminars/conferences relevant to duties	Quality
6.8 Reports and disseminates to other instructors in their department/faculty	Quantity
6.9 Implementation of the knowledge gained from training/seminars/conferences	Quality

2.3 The Development of Indicators and Weighting

When developing an indicator form a definition of a concept to be measured, there are 3 steps as follows: (Johnstone, 1981)

1) select the component variable
2) determine the method of combination
3) define the weights to be applied to each variable

2.3.1 Selecting the component variables

The selection of variables for combination to form a theoretically defined indicator can be effected in a variety of ways. Normally, variables that might be included can be easily identified. Making a selection from just these variables, however, can lead to a bias in the underlying construct measured by the final form of the indicator. A more systematic approach than this is required.

To ensure that all relevant variables are considered when a selection is made to define an indicator, it is necessary firstly to develop a statement concerning the sorts of characteristics the indicator is to measure. In other situations an expert consensus must decide exactly what characteristics of an education system are to be covered by a defined indicator. In arriving at such a consensus, a two or three dimensional grid delineating the main factors involved in the characteristic can sometimes be a useful aid. If a definition indicates that too many variables might have to be combined to form a single indicator, a very complex concept is indicated as having to be measured. In such a situation, it may be preferable to reduce this complexity so as to avoid confounding different aspects of the concept.

Once the various characteristics to be included have been identified variables can be selected or defined to measure the specific features of each characteristic. Multiple measurements of a single characteristic are not permissible if parsimonious definition is to be achieved. Also, variables suspected to contain large errors of measurement should not be selected for combination. It might even be better to defer temporarily the development of an indicator than to proceed with the formation of a distorted composite.

When a choice still exists among a number of variables after considerations such as those outlined above, selection can sometimes be assisted by inter-correlating different variables to determine the extent of their empirical overlap. Diagram 2.1 demonstrates a possible situation, which could arise when selecting variables to combine into a composite index. There, variable 1 and 2 are represented as being highly interrelated to the concept being measured. Such a determination would have to be ascertained by judgment rather than by empirical

means when the indicator has not yet been formed. Both of these variables are also highly inter-related. As they are probably measuring a similar characteristic, they should not be both included in the variable set. Variable 3 has a low correlation with both variables 1 and 2 and a moderate correlation with concept. Therefore, it contributes to the definition in a more independent way. Combining variable 1 or 2 with variable 3 in this case would best form the indicator of the concept. In many situations, this empirical check on the behavior of variables cannot be employed. Then only expert opinion, guided by the purpose for which the indicator is being developed, can be used.

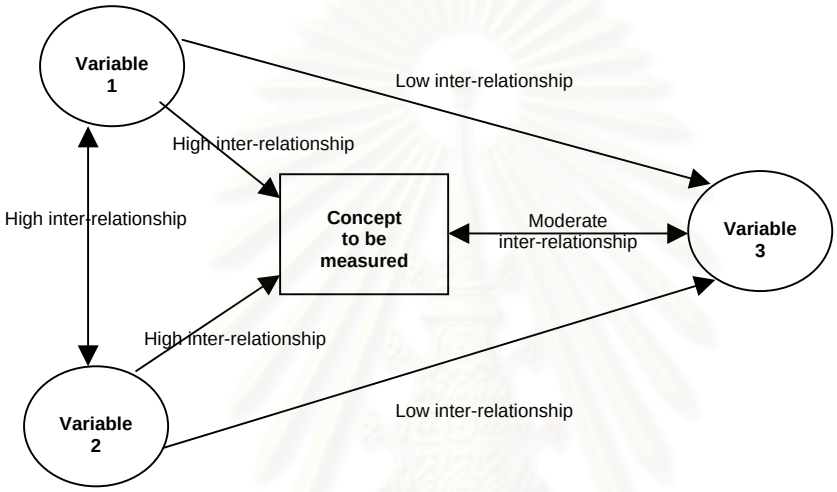


Diagram 2.1 Hypothetical inter-relationships among three variables and a concept to be measured (Johnstone, 1981: 73)

2.3.2 Determining the method of combination

The vast majority of theoretically defined indicators are formed by simply adding together the component variables. Sometimes this is the most correct method of combination. At other times, it is not correct for addition does not always reflect the combination of variables as defined by the concept for the indicator. Multiplication may be a more appropriate technique of combination.

When addition is the method used for combining variables, an implicit assumption is made that any one variable can substitute or compensate for another. For example, an indicator (I) might be formed from two component variables (V_1 and V_2) as follows:

$$I = V_1 + V_2 \qquad (2.1)$$

An average total value could be calculated for this indicator either by both component variables having average values or by one variable being considerably above the mean and the other being considerably below the mean.

Multiplying the values for the component variables together to form an indicator implies a different concept from that of addition. An indicator formed in this way can be represented as:

$$I = V_1 \times V_2 \qquad (2.2)$$

Under this method of combination, compensation for a small value for one variable cannot occur as readily. The operation implies instead that a high indicator value can only be calculated when all component variables have high values. The value of one variable thus builds upon that of another.

One implication of the difference between the two methods of combination is the way in which an average value can be formed. When the operation of addition is used, an arithmetic mean can define the indicator value. That is:

$$I = \frac{V_1 + V_2 + V_3 + ... + V_n}{n} \qquad (2.3)$$

or for weighted variables

$$I = \frac{w_1V_1 + w_2V_2 + w_3V_3 + ... + w_nV_n}{w_i} \qquad (2.4)$$

where w_i is the weight for the i^{th} variable V_i.

When multiplication is used, the geometric mean is appropriate. In its unweighted form, this can be expressed as

$$I = \sqrt[n]{V_1.V_2.V_3.....V_n} \qquad (2.5)$$

and in its weighted form

$$I = \sqrt[n]{V_1^{W_1}.V_2^{W_2}.V_3^{W_3}.....V_n^{W_n}} \qquad (2.6)$$

The computation of a value form this equation is facilitated by the use of logarithms.

2.3.3 Defining the weights

The distinction between theoretically defining an indicator rather than allowing empirical definition is most marked when the problem of determining weights for the component variables arises. Partly because of the subjective nature of the decisions, which have to be made, the definition of weights in this method of developing indicators can be a most contentious area. Its importance cannot be diminished because the selection of one set of weights rather than another can substantially change any results.

The most common approach used to select weights is to define all component variables as being equal in value and importance. Hence each variable is assigned unit weight. If this could always be done, no weighting problem would exist. The

alternative to this approach is to use differential weights for the variables. Here planners and researchers nominate their own values for the weights – with or without justification.

Expert judgment from sources other than the researcher can also determine weights. Sometimes this might be very simplistic and merely involve estimating percentages responding in particular categories of a questionnaire. In other situations, it can be very much more sophisticated.

A second approach to defining weights is to measure the effort required to establish the value of a particular variable. One such measure of effort is the time taken or allocated to produce a variable value. This approach assumes that if the time taken to do something is longer for on variable that for another, then that variable should be weighted more (or less) heavily.

A third method of defining weights is to use agreed standards. In education, few opportunities overtly present themselves to allow the identification of standards.

The selection of one method of determining the weight for theoretically defined indicators rather than another depends on many considerations. The nature of the indicator itself must be taken into account; so too must the nature of the variables being combined. No definite rules can therefore be proposed to guide indicator development of this type.

2.4 Test of Linearity and Non-linearity

Analytic Geometry Theory has been used to consider linearity of scores as follows: (Fuller, 1967: 28-29)

The first-degree equation or the linear equation is

$$Ax + By + C = 0 \qquad (2.7)$$

Where A, B, C are constants with A and B not both zero, is a general equation of the first degree. It can be proven that the locus, or graph, of this equation is a straight line by showing that all points of the locus lie on a line and that the coordinates of all points of the line satisfy the equation.

Let P_1 (x_1, y_1) and P_2 (x_2, y_2) be two points of the graph (Figure 2.2). Then the coordinates of these points satisfy Equation 2.7, and therefore we have

$$Ax_1 + By_1 + C = 0 , \qquad (2.8)$$

$$Ax_2 + By_2 + C = 0 . \qquad (2.9)$$

By subtraction, these equations yield

$$A(x_1 - x_2) + B(y_1 - y_2) = 0, \qquad (2.10)$$

$$B(y_1 - y_2) = -A(x_1 - x_2), \qquad (2.11)$$

And if $B \neq 0$

$$\frac{y_1 - y_2}{x_1 - x_2} = -\frac{A}{B} \qquad (2.12)$$

Equation 2.12 shows that the slope of a line passing through two points of the graph is – (A/B). Therefore if P_3 (x_3, y_3) is any other point of the locus, the slope of the segment P_1P_3 is also – (A/B). From the equality of these slopes, we conclude that P_1, P_2 and P_3, and hence all points of the locus, lie on a line.

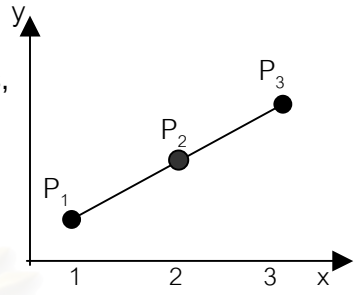


Figure 2.2

To conclude, Let P_1 (x_1, y_1), P_2 (x_2, y_2) and P_3 (x_3, y_3) are the coordinate of year (x) and performance score (y) and Let slope of P_1P_2 is m_1 and slope of P_1P_3 is m_2. The point P_1, P_2 and P_3 will be lie on the line (linear) if $m_1=m_2$. (where $m_1=\frac{y_1 - y_2}{x_1 - x_2}$ and $m_2 = \frac{y_1 - y_3}{x_1 - x_3}$)

2.5 Measuring Change

For methodology to measure change, the International Encyclopedia of Educational Evaluation (1990) divided it into 2 main methodologies that are 'measuring change with 2 waves of data' and 'measuring change with multi-wave data' (Willet, 1990).

2.5.1 Measuring change with 2 waves of data

Measuring change with 2 waves is used for calculating individual change i.e. difference change or gain score – difference of pre and post measurement, Webster and Bereiter's (1963) reliability-weight measure of change, Lord's (1956) regression-based estimated true change, Residual change scores – obtained by estimating residuals from the population regression of true final status on true initial status (Willet, 1990). It is true pre/post (or two-wave) measurement is the most common design in the study of change, and two repeated observations do indeed constitute a longitudinal study. However, two observations are not adequate for studying the form of change. Two observations can only estimate the amount of change in a straight line. Rogasa showed that If the rate of growth is not constant, but depends on time, the amount of change will depend crucially on the times of measurement, and observations of individuals at a difference set of two time points may give contradictory results (Gottman, 1995).

2.5.2 Measuring change with multiwave data

In the last decade, researchers have moved beyond the limitation of the pre/post or two-wave measurement. They prefer and widely use the methodology to measure change with multi-wave data. The methodology to measure change for individuals can be classified into 4 groups: Structural Equation Model, Adopting Dynamic Theory, Applying Mathematics Principles, and Hierarchical Linear Model.

Structural equation model as it is in Latent Curve Analysis by Meredith and Tisak (1990), which aims mainly on statistics hypothesis testing about the relation between observed variables and unobserved variables. Data analysis for Measuring change is on a condition that the number of respondents has to be big at least 300 people (Bijlevled and et a. 1998) which may cause problems for a university which has a small group of instructors.

Measuring change adopting dynamic theory as in Multidimensional Latent Trait Model by Embertson (1991) to measure learning change. The results obtained from this model were the amount of individual change and quality of tests based on the basic assumption that examiners used their cumulative ability. It is also in Gultman Simplex Model by Collins and Cliff (1988). In this model, individual change and group change were found on the assumption that cumulative ability was used in the measurement, which is only on one direction and never decreased. If not, the measurements for research would not satisfy the assumptions, and some variables e.g. instructor performance may have many patterns of change.

From the above concept, Structural Equation Modeling has problems for calculating performance score in Saint John's University because it has a small group of instructors (60 instructors meeting determined criteria), and also measuring change adopting dynamic theory would not be satisfied because the variables in this research (performance of the instructor) may have many patterns of change (not only one direction). Thus, both concepts will not be described in this research.

The other methods, which are measuring change applying Mathematics principles and Hierarchical Linear Model, will be described and are separated into 2 groups as follows:

Measuring methods for calculating linear change scores

Measuring methods for calculating non-linear change scores

2.5.2.1 Measuring Methods for Calculating Linear Change Scores

1) Slope

In general, movement along a straight line of slope m results in m vertical units gained for each horizontal unit gained. Thus, the slope, m, is the rate of change of vertical position (y) with respect to horizontal position (x), as is illustrated below (Figure 2.3) (Piascik, 1994:15).

If (x_1, y_1) and (x_2, y_2) represent two points through which a straight line passes and we move along the straight line from (x_1, y_1) to (x_2, y_2) then the vertical change is denoted by Δy and our horizontal change is denoted by Δx.

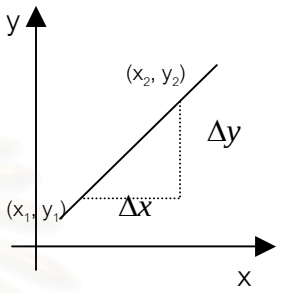


Figure 2.3

Furthermore, $\Delta y = y_2 - y_1$ and Δx $y_2 - y_1$

And the slope m of the straight line is determined by the following formula (Piascik, 1994: 13).

$$\text{Slope} = m = \frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1} \qquad (2.13)$$

2) Regression

The general form of the regression equation is (Lind and Mason, 1997: 357)

$$\hat{Y} = a + bX \qquad (2.14)$$

where $\hat{Y}$ is the predicted value of the Y variable for a selected X value

X is any value of independent variable that is selected

a is the Y-intercept. It is the estimated value of Y when X=0

b is the slope of the line, or the average change in $\hat{Y}$ of one unit (either increase or decrease) in the independent variable X

To obtain the optimal values of a and b, we will begin with the expression to be minimized $\left[\sum (Y - \hat{Y})^2\right]$. Substituting bX + a for $\hat{Y}$ as (Howell, 1992: 332-333)

$$\sum (Y - \hat{Y})^2 = \sum (Y - bX - a)^2 \qquad (2.15)$$

$$= \sum (Y^2 + b^2 X^2 - a^2 - 2aY - 2bXY + 2abX) \qquad (2.16)$$

The derivative with respect to a is

$$\frac{d(Y - \hat{Y})^2}{da} = \sum (2a - 2Y + 2bX) \qquad (2.17)$$

$$= \quad 2Na - 2\sum Y + 2b\sum x \qquad (2.18)$$

Setting this equal to zero,

$$2Na - 2\sum Y + 2b\sum x \;=\; 0 \qquad (2.19)$$

$$Na + b\sum X \;=\; \sum Y \qquad (2.20)$$

The derivative with respect to a is

$$\frac{d(Y-\hat{Y})^2}{db} \;=\; \sum(2bX^2 - 2XY + 2aX) \qquad (2.21)$$

$$= \quad 2b\sum X^2 - 2\sum XY + 2a\sum X \qquad (2.22)$$

Setting this equal to zero,

$$2b\sum X^2 - 2\sum XY + 2a\sum X \;=\; 0 \qquad (2.23)$$

$$b\sum X^2 - a\sum X \;=\; \sum XY \qquad (2.24)$$

The derived equations are

$$Na + b\sum X \;=\; \sum Y \qquad (2.25)$$

and

$$b\sum X^2 - a\sum X \;=\; \sum XY \qquad (2.26)$$

From equations 2.25 and 2.26, It is a simple matter to solve for a and b. Solving for a as

$$a = \frac{\sum Y - b\sum X}{N} \qquad (2.27)$$

$$a = \overline{Y} - b\overline{X} \qquad (2.28)$$

From equations 2.25 and 2.26, to solve for b as

$$b = \frac{N\sum XY - \sum X\sum Y}{N\sum X^2 - (\sum X)^2} \qquad (2.29)$$

where b in Equation 2.29 is the average change in $\hat{Y}$ of one unit (either increase or decrease) in the independent variable X

3) Time Series

The linear trend equation can be written as (Gupta, 1993: 309)

$$Y_c \;=\; a + bX \qquad (2.30)$$

Where Y_c is the calculated trend value

X is a number of time

a represents the intercept on Y-axis

b represents the slope of the line, i.e., the amount of change in Y for a

unit change in the value of X

The method consists of choosing the values of a and b such that the sum of the squared deviations $(Y - Y_c)^2 = (Y - a - bX)^2$ is the least.

For this, use of the following normal equations. (Gupta, 1993: 310)

$$Na + b\sum X = \sum Y \qquad (2.31)$$

and

$$b\sum X^2 - a\sum X = \sum XY \qquad (2.32)$$

From equations 2.31 and 2.32, It is a simple matter to solve for a and b as same as regression method. The derived equations are

$$a = \overline{Y} - b\overline{X} \qquad (2.33)$$

and

$$b = \frac{N\sum XY - \sum X \sum Y}{N\sum X^2 - (\sum X)^2} \qquad (2.34)$$

where b in equation 2.34 represents the amount of change in Y for a unit change in the value of X

4) Hierarchical Linear Model (Application of Linear Growth Model)

Hierarchical Linear Model (application of linear growth model) was developed by Bryk & Raudenbush (1987). This methodology was considered to be a good methodology for measuring change (Bijileveld *et al*, 1998; Kamlan, 1997; Rer-juntig, 1999) because it composed many statistics principles for calculating change and it can examine the reliability of initial status and change and investigate correlates of status and individual change (Bryk and Raudenbush, 1987).

In many situations, particularly when the number of observations per individual are few (e.g. three or four occasions), It is convenient to employ a linear individual growth model.

The level 1 model is (Bryk and Raudenbush, 1992: 135)

$$Y_{it} = \pi_{0i} + \pi_{1i} a_{it} + e_{it} \qquad (2.35)$$

They assume that Y_{it}, the observed status at time t for individual i, is a function of a systematic growth trajectory or growth curve plus random error. a_{it} is the age at time t or the time t for subject i, π_{0i} and π_{1i} are the growth trajectory parameter for subject i. And the errors e_{it} are assumed to be independent and normally distributed with common variance σ^2

A Random-Coefficient Regression Model

Equation 2.35 specifies the Level 1 model. At Level 2, we begin with the simplest person level model: (Bryk and Raudenbush,1992: 175-176)

$$\pi_{0i} = \beta_{00} + U_{0i} \quad (2.36)$$

$$\pi_{1i} = \beta_{10} + U_{1i} \quad (2.37)$$

Here the intercept, π_{0i} and the linear rate of change, π_{1i}, determine the change for each subject. The random effects (U_{0i}, U_{1i}) are assumed to be bivariate normal with zero means, variance τ_{00}, τ_{11}, and covariance τ_{10}. The regression coefficient β_{00} and β_{10} are termed fixed effects. Equation 2.36 and 2.37 are simple Level 2 model (an unconditional model) in that no Level 2 predictors for either π_{0i} or π_{1i}. This model provides useful empirical evidence for determining a proper specification of the individual growth equation and baseline statistics for evaluating subsequent Level 2 models. When equations 2.35, 2.36 and 2.37 combined, the resulting equation may be written (Raudenbush, 1995)

$$Y_{it} = \beta_{00} + \beta_{10}\, a_{it} + \in_{it} \quad (2.38)$$

Where $\in_{it} = U_{0i} + U_{1i}\, a_{it} + e_{it}$. Two aspects of the error $\in_{it}$ render Equation 2.38 inappropriate for estimation via ordinary least-square regression. First, these errors are correlated within subjects by virtue of the fact that every time-series observation for subject i shares random effect U_{0i} and U_{1i}. Second, the errors are heteroscedastic, given the dependence of their variance on a_{it}. Under these conditions, and with balanced data, ordinary least-squares estimates of the fixed effect β_{00} and β_{10} would be efficient in the case of Equation 2.38, but the standard error estimates would not. When the data are unbalanced and/or when time-varying covariates are added to the model, ordinary least-squares regression would yield inefficient estimates of the fixed effects as well. To solve these problems, a variety of algorithms can be employed the produce maximum-likelihood estimates of all the parameters of Equation 2.38, including the variance and covariances. HLM program of Raudenbush, Bryk and Congdon (2000) were utilized to compute such estimates.

Empirical Bayes (''EB'') estimates of randomly varying level-1 coefficients for each unit i are optimal composites of an estimate based on the data from that unit and an estimate based on data from other similar units. Intuitively, we are borrowing strength from all of the information present in the ensemble of data to improve the

level-1 coefficient estimates for each of the i units. These "EB" estimates are also referred to as "shrunken estimates" of the level-1 coefficients. They are produced by HLM as part of the residual file output (Raudenbush, Bryk and Congdon, 2000)

Empirical Bayes provides a composite estimator and depends on the reliability of the ordinary least square estimate. When this is highly reliable, the HLM estimate for an individual's growth rate will lean heavily on the individual time series data. (Bryk and Raudenbush, 1987) Moreover, Empirical research showed that the composite estimator has a smaller mean squared error than the least square estimator. (Efron & Morris, 1979; Morris, 1983 and Raudenbush, 1995)

2.5.2.2 Measuring Methods for Calculating Non-Linear Change Scores

1) Second-Order polynomial regression and Rate of change

The general form of the second order polynomial regression is (Neter and Wasserman, 1974: 274)

$$\hat{Y} = a + bX + cX^2 \qquad (2.39)$$

where $\hat{Y}$ is the predicted value of the Y variable for a selected X value

X is any value of independent variable that is selected

a is the mean response of Y when X=0 (a has no separate meaning of its own in the model)

b is the linear effect coefficient

c is the curvature effect coefficient

To obtain the values of unknown constant; a, b, c in Equation 2.39, the equation 2.40 - 2.42 that obtained from least square method were used as follows: (Gupta, 1993: 325)

$$\text{I} \quad na + b\Sigma(X) + c\Sigma(X^2) = \Sigma(Y) \qquad (2.40)$$

$$\text{II} \quad a\Sigma(X) + b\Sigma(X^2) + c\Sigma(X^3) = \Sigma(XY) \qquad (2.41)$$

$$\text{III} \quad a\Sigma(X^2) + b\Sigma(X^3) + c\Sigma(X^4) = \Sigma(X^2Y) \qquad (2.42)$$

(where n = number of years)

Set X have values –1, 0, 1 respectively (for 3 year periods). Thus, $\Sigma(X) = 0$ and $\Sigma(X^3) = 0$ and equation 2.40 to 2.42 would be follows:

$$\text{I} \quad na + c\Sigma(X^2) = \Sigma(Y) \qquad (2.43)$$

$$\text{II} \quad b\Sigma(X^2) = \Sigma(XY) \qquad (2.44)$$

$$\text{III} \quad a\Sigma(X^2) + c\Sigma(X^4) = \Sigma(X^2Y) \qquad (2.45)$$

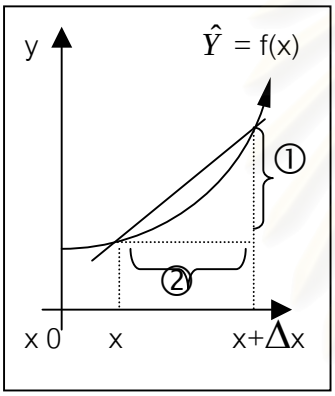
or $$a = \frac{\sum(Y) - c\sum(X^2)}{n} \quad (2.46)$$

$$b = \frac{\sum(XY)}{\sum(X^2)} \quad (2.47)$$

$$c = \frac{n\sum(X^2Y) - \sum(X^2)\sum(Y)}{n\sum(X^4) - (\sum X^2)^2} \quad (2.48)$$

From Equation 2.39, $\hat{Y} = f(x)$, it have 2 methods for calculating rate of change, which are average rate of change and instantaneous rate of change.

1) **Average rate of change** of a function $f(x)$ is given by the formula (Piascik, 1994: 119)



$$\frac{\Delta y}{\Delta x} = \frac{f(x+\Delta x) - f(x)}{\Delta x} \quad (\Delta x \neq 0) \quad (2.49)$$

The above expression, called the difference quotient, gives the slope of the secant line passing through (x, $f(x)$) and (x+Δx , $f(x+\Delta x)$), as illustrated in Figure 2.4.

Figure 2.4 ① $\Delta y = f(x+\Delta x) - f(x)$ ② Δx

2) **Instantaneous rate of change** of $\hat{Y}$ with respect to x at any point (x, $f(x)$) on a function $\hat{Y} = f(x)$ is given by the derivative $f'(x)$ which is defined by (Piascik, 1994: 129)

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} \quad (2.50)$$

$$\therefore \quad f'(x) = b + 2cx \quad (2.51)$$

Provided this limit exits. Graphically, the derivative is the slope of the straight line tangent to the graph of the function at (x, $f(x)$)

2) Hierarchical Linear Model (Application of Quadratic Growth Model)

This method is used for measuring change scores of the instructors who have quadratic change. The model at Level 1 is now of the form: (Bryk and Raudenbush,1992 : 140)

$$Y_{it} = \pi_{0i} + \pi_{1i}(a_{it} - L) + \pi_{2i}(a_{it} - L)^2 + R_{it} \quad (2.52)$$

The specific or priori centering constant, L, for the Level 1 predictors that are powers of a_{it}. Each of the growth parameters in Equation 2.52 has a substantive meaning. The intercept, π_{0i}, represents the status of subject i at time L. The linear component, π_{1i}, is the instantaneous rate of change for subject i at time L, and π_{2i} captures the curvature or acceleration in each growth trajectory, the status and instantaneous rate parameters depend on the particular choice of value for L.

The level 2 model

Equation 2.52 characterizes each subject 's trajectory form academic year 1998 to 2000 by three parameters: π_{0i}, π_{1i} and π_{2i}. The formula of a simple level 2 model (unconditional model) are

$$\pi_{0i} = \beta_{00} + U_{oi} \quad (2.53)$$

$$\pi_{1i} = \beta_{10} + U_{1i} \quad (2.54)$$

$$\pi_{2i} = \beta_{20} + U_{2i} \quad (2.55)$$

The regression coefficients β_{00}, β_{10} and β_{20} are termed fixed effects. The random effects (U_{0i}, U_{1i}) are assumed normal with zero means, variance and covariance as

$$T = \begin{pmatrix} \tau_{00} \\ \tau_{10}\ \tau_{11} \\ \tau_{20}\ \tau_{21} \end{pmatrix} = \begin{pmatrix} \mathrm{Var}(\pi_{0i}) \\ \mathrm{Cov}(\pi_{1i}, \pi_{0i})\ \ \mathrm{Var}(\pi_{1i}) \\ \mathrm{Cov}(\pi_{2i}, \pi_{0i})\ \mathrm{Cov}(\pi_{2i}, \pi_{1i}) \end{pmatrix}$$

In general, the growth rate at any particular time is the first derivative of the growth model evaluated at that time. For quadratic growth,

Growth rate at time t = $\pi_{1i} + 2\pi_{2i}(a_{it} - L)$ (2.56)

2.5.3 Research Related to Measuring Change

Chan and et al. (2000) used multivariate latent growth modeling to conceptualize and analyze intraindividual changes in children's social skills and interindividual differences in these changes in home and social settings. Parent and teacher ratings assessing children's social skills at home and at school settings, respectively, were obtained for a sample of 378 children at 4 time points spaced at approximately 12-month intervals over a 4-year period from Kindergarten to Grade 3. Results showed that, for the initial status at Kindergarten, there were significant individual differences in social skills in both home and school settings and

a significant positive association between initial status in social skills in the two settings. Systematic between-settings differences in children's social skill development were found. Social skills development at home was best described with a nonlinear trajectory in which skills increased from Kindergarten to Grade 2 with a substantially larger increase from Grade 1 to Grade 2 than from Kindergarten to Grade 1, and then remained relatively constant from Grade 2 to Grade 3. In contrast, social skills development at school was best described with a negative linear trajectory in which skills decreased at a constant rate from Kindergarten to Grade 3. The differences in social skills development may derive from the fact that different teachers with different expectations regarding social skills provided ratings each year while the same parent was the source of at-home social skills ratings. There were significant individual differences in growth rates in the school as well as the home setting. Evidence of between-settings differences in social skills development were obtained from differential patterns of associations between growth parameters (initial status and growth rate) and individual predictors (family income, parent education, child verbal skills) across settings.

MacCallum and et al. (1997) illustrated the multilevel linear model of change using data from a study of physiological response to marital conflict in older couples. Thirty-one married couples between the ages of 55 and 75 participated in this study. The average length of marriage for this sample was approximately 42 years. The couples were admitted to the Ohio State University Clinical Research Center at 7 a.m. and remained there for an 8-hour period during which a variety of activities occurred. Upon admission, a heparin well was inserted in each subject's arm to allow for easy and unobtrusive drawing of blood during subsequent activities. The present analyses make use of data from 5 blood samples at 5 time points. From each blood sample, endocrine assays were conducted to measure levels of 3 hormones: cortisol, norepinephrine, and adrenocorticotropic hormone (ACTH). Results showed that, there is a clear linear trend in norepinephrine, with the mean level increasing over time, but there is no systematic trend in the means of the other two hormones. However, the considerable overlap in confidence intervals across time for ACTH and cortisol indicates no strong evidence for systematic nonlinear change in the means. Moreover, Many multilevel models have equivalent representations as latent curve models (LC), although some multilevel models that are nonlinear in their parameters appear not to have corresponding LC representations. With regard to the study of variables that may serve as correlates, predictors, or consequences of change, the LC approach is more flexible than currently available multilevel methods. The

conventional multilevel framework is more restricted in the type of associations that can be investigated regarding predictors of change.

Silverstein and Long (1998) analyzed data from the longitudinal study of generations are used to identify patterns of change in grandparents' perceptions of affection and in-person contact and geographic proximity with adult grandchildren over five points of measurement between 1971 and 1994. Hierarchical linear modeling reveals quadratic trends in both growth curves. Affection declines over the first 14 years and modestly reverses. When cohorts are equated on age, later cohorts of grandparents decline more rapidly in contact and proximity, suggesting that the grandparent role has changed in recent history.

Shin and et al (2000) examined the technical adequacy of curriculum-based measurement (CBM) for assessing student growth over time. Participants were 43-second graders whose reading performance was measured monthly over 1 school year with the maze task. Technical characteristics of the CBM maze task were examined in terms of reliability, sensitivity, and validity for assessing student growth using hierarchical linear models. Results showed that the maze task had good alternate-from reliability, with a mean coefficient of .81 with 1-to 3-month intervals between testing. The maze task also sensitively reflected improvement of student performance over a school year and revealed interindividual differences in growth rates. Finally, growth rates estimated on repeated maze scores were positively related to later reading performance on a standardized reading test; in addition, although a significant difference was not found, general education students appeared to develop reading proficiency faster than remedial education students. Results support the use of the maze task as a reliable, sensitive, and valid data collecting procedure for assessing reading growth.

Chou, Bentler and Pentz (1998) compared two statistical approaches for modeling growth across time. The two statistical approaches are the multilevel model (MLM) and latent curve analysis (LCA), which have been proposed to depict change or growth adequately. These two approaches were compared in terms of the estimation of growth profiles represented by the parameters of initial status and the rate of growth. A longitudinal data set obtained from a school-based substance-use prevention trial for adolescents was used to illustrate the similarities and differences between the two approaches. The results indicated that the two approaches yielded very compatible results. The parameter estimates associated with regression weights are the same, whereas those associated with variances and covariances are similar. The MLM approach is easier for model specification and is more efficient

computationally in yielding results. The LCA approach, however, has the advantage of providing model evaluation, that is, an overall test of goodness of fit, and is more flexible in modeling and hypothesis testing.

Williamson, Appelbaum and Epanchin (1991) studied longitudinal analyses of academic achievement. The data consist of 8 waves of achievement scores collected in the spring of each year from 1978 to 1985 and 3 ability scores collected in the fall of 1979 for a cohort of students as they progressed from Grade 1 to Grade 8. From the database, 529 individuals with complete Reading Total records and 527 individuals with complete Mathematics Total records were used in the study. Results showed that, Only 38 individuals (7.2%) showed a significant quadratic component for Reading and 26 individuals (4.9%) showed a significant quadratic component for Mathematics. The straight-line model is assumed to be appropriate for all students. Longitudinal data analysis are feasible for school district when a suitable longitudinal database has been maintained.

Khamlan (1997) developed a growth curve model for longitudinal analysis of change in English vocabulary achievement, and to analyze the change in English vocabulary achievement with application of HLM program. The sample consisted of 603 Grade 6 students at 8 time points spaced at approximately 2-week intervals over a 4-month period. Results showed that, the quadratic growth model could explain more variance in change of English vocabulary achievement than linear growth model. The student had a growth rate in learning English vocabulary of an average .38 words per 2 weeks, and the mean of acceleration was -0.02. The estimated reliabilities for initial status, growth rate and acceleration were .99, .58 and .43 respectively. The initial status had a positive relationship with growth rate and acceleration. The intelligence and motivation significantly affected the initial status, attitude and dummy variable female significantly affected the acceleration.

Ruachantuk (1999) studied the analysis results of change in environmental knowledge obtained from longitudinal analysis by HLM and LISREL models. The sample consisted of 509 Grade 8 students at 7 time points in 4 months. Results showed that, there is a linear change in environmental knowledge, with the mean level increasing over time. The initial status had a positive relationship with rate of environmental knowledge change and the environment attitude were significantly affected the rate of change. Moreover, the longitudinal analysis of change by LISREL model was more explicable the longitudinal analysis of change that HLM model because LISREL model had a lower error, but HLM is easier for model specification

and is more efficient computationally in yielding results, especially individual initial status and rate of change.

Wijitwanna (2000) compared the efficiency of latent growth curve model, multilevel models and quasi-simplex models in measuring univariate and multivariate longitudinal change in mathematical abilities: calculation and problem solving. The quasi-simplex models with latent growth was developed by Wijitwanna in order to explain true initial and change parameters. The 5 wave data were collected from the population of 469 Grade 8 students in Samusongkram Province in the 1998 academic year. Results showed that there is a linear trend in calculation, with the mean level increasing over time, but there is a quadratic trend in the mean of problem solving. Latent growth curve models had the best efficiency in measuring univariate and multivariate longitudinal change.

Tangsakulruanglai (1998) compared the efficiency among four latent variable growth curve models in studying longitudinal changes in Mathematics achievement and physical development with application of LISREL models. The sample consisted of 406 Grade 6 students in schools under the jurisdiction of Bangkok metropolis and 592 Grade 7 students in the schools under the office of Pitsanulok primary education at 5 time points. Results showed that, the growth curve of the students' Mathematics achievement was a downward quadratic while the curves of the students' weight and height were upward quadratic. The result from an application of four latent variable growth curve models indicated that LISREL models in the form of the latent growth curve model with free parameter and unequal disturbance variance was the most efficient one in studying longitudinal changes in development of Mathematics achievement and physical development on weight and height.

2.5.4 The development of formulae for measuring change

From the above methodology for measuring change, measuring change by applying mathematics principles such as Regression, Slope, Time series, or Polynomial are simple methods that are easy to calculate change score. But they cannot find the difference of change scores over a 3 year period in instructors who have different patterns of change, if they have the same scores in the first and third year of the measurement. The other methods such as Structural Equation Model, measuring change adopting Dynamic Theory and Hierarchical Linear Model are complex for general people (because of advance methodology e.g. advance statistics, measurement theory) and it is difficult to interpret the meaning of change scores and they also usually conducted with a large sample size.

Therefore, it is necessary to have new formulae that are easy to calculate and translate the meaning of change scores. Furthermore, the formula could calculate change scores for individuals without impact of sample size.

In this research, the academic years during the introduction of quality assurance in TU and SU are 1998, 1999 and 2000. The instructor's performance scores over a 3 year period may have linear and non-linear change. Thus, separate formulae for calculating linear and non-linear change scores were developed. The Law of Initial Values (LIV) and Floor and Ceiling Effects with the following concepts were considered when developing the formulae.

"According to the LIV, as scores approach a floor or a ceiling, the same external stimulus should produce smaller changes. Thus the amount of change would be a function of the starting level (i.e., the initial difference from the floor or ceiling). When approaching a ceiling, the higher scores will show a smaller increment, which will result in a negative for r_{dx} (d is change, x is initial value). When approaching a floor, the smaller scores decrease less, which also results in negative r_{dx}." (Jamieson, 1995: 40)

The above concept implies that the instructor's performance change score depends on instructor's performance especially in the initial year. If an instructor has a high performance score in the initial year; it will be hard for him to increase his score. It shows that the increment of performance score was limited by the ceiling effect. In addition, for an instructor who had a low performance score in the initial year, his score decreased less than an instructor who had a higher score. It shows that the decrease in performance score was limited by the floor effect. Thus, the performance change score is a function of change of observed scores, initial score, and floor and ceiling effects.

Moreover, there are many patterns of non-linear change (i.e. parabola, quadratic, exponential etc.) in measuring non-linear change. From the concept of organization theory that Sudchari *et al*. (1999) developed from Schewiger *et al*. (1986), that is "Organization survives and grows when it has stability and high adaptability." Instructors are a vital part of any university and should improve their performance regularly. Thus, the pattern that should have the highest positive change is the pattern that performance scores increase regularly; and the pattern that should have the highest negative change is the pattern that performance scores decrease regularly. This concept was considered when developing formulae for calculating change scores.

In this research, therefore, the formulae for calculating linear or non-linear change scores of an instructor's performance were developed with consideration of the law of initial values, floor and ceiling effects and organization theory.

The method of Bryk & Raudenbush (1987), Hierarchical Linear Model, was used as a validation criterion in this research. It was selected because it is a good method for measuring change (Bijileveld *et al*, 1998; Khamlan, 1997; Ruachantuk, 1999), which composed many statistics principles for calculating change (Bryk and Raudenbush, 1987). But it has a limitation in estimating parameters in the case of respondents less than 100 (Bijileveld and et al, 1998) and it is complex for general people because of advanced methodology.

The two Hierarchical Linear Models of Bryk & Raudenbush (1987) that are possible for calculating change score in this research are Hierarchical Linear Model (application of linear growth model) for linear change and Hierarchical Linear Model (application of quadratic growth model). The principles of them for calculating performance change score are as follows:

1) Hierarchical Linear Model (Application of Linear Growth Model)

This method is used for measuring change scores of the instructors who have a linear change. According to a simple linear model for individual change, there is a tendency for the performance score of each instructor to change at steady rate from academic year 1998 to academic year 2000. The academic year variable is a deviate from academic year 1998 according to the following table:

Academic Year (AY)	1998	1999	2000
AY – 1998	0	1	2

The level 1 model is

$$Y_{it} = \pi_{0i} + \pi_{1i}(AY - 1998)_{it} + e_{it} \qquad (2.57)$$

where Y_{it} = instructor's performance score for instructor i at time t, i = 1,2,…,238 for TU instructors or i = 1, 2,…,52 for SU instructors, t = 1, 2, 3

$(AY\text{-}1998)_{it}$ = academic year of instructor i at time t minus 1998 so that $(AY\text{-}1998)_{it}$ is 0, 1, 2 at academic years 1998, 1999 and 2000

π_{0i} = intercept of instructor i, so that given the coding of (AY-1998), π_{0i} is the expected performance score of instructor i at academic year 1998

π_{1i} = the expected rate of change per academic year in the performance score for instructor i

e_{it} = the random within-instructor error of prediction for instructor i at time t, conditional on that instructor's chance parameters π_{0i} and π_{1i}

The level 2 model

Equation 2.57 characterizes each instructor 's trajectory form academic year 1998 to 2000 by two parameters: π_{0i} and π_{1i}. The formula of a simple level 2 model (unconditional model) as

$$\pi_{0i} = \beta_{00} + U_{0i} \qquad (2.58)$$

$$\pi_{1i} = \beta_{10} + U_{1i} \qquad (2.59)$$

where π_{0i} = intercept of instructor i, so that given the coding of (AY-1998), π_{0i} is the expected performance score of instructor i at academic year 1998

π_{1i} = the expected rate of change per academic year in the performance score for instructor i

β_{00} = grand mean performance score at academic year 1998

β_{10} = grand mean rate of change in performance score

U_{0i} = random effect of instructor i on the performance score at academic year 1998

U_{1i} = random effect of instructor i on the rate of change in performance score

To conclude, the performance score change of instructor i who has linear change over a 3 year period were calculate by using Equation 2.59 ($\pi_{1i} = \beta_{10} + U_{1i}$).

2) Hierarchical Linear Model (Application of Quadratic Growth Model)

This method used for measuring change scores of the instructors who have curvilinear change. The model at Level 1 is now of the form:

$$Y_{it} = \pi_{0i} + \pi_{1i}(AY-1998)_{it} + \pi_{2i}(AY-1998)_{it}^{2} + e_{it} \qquad (2.60)$$

where Y_{it} = instructor's performance score for instructor i at time t, i = 1,2,…,238 for TU instructors or i = 1, 2,…,52 for SU instructors, t = 1, 2, 3

$(AY-1998)_{it}$ = academic year of instructor i at time t minus 1998 so that $(AY-1998)_{it}$ is 0, 1, 2 at academic years 1998, 1999 and 2000

π_{0i} = intercept of instructor i, so that given the coding of (AY-1998), π_{oi} is the expected performance score of instructor i at academic year 1998

π_{1i} = the instantaneous rate of change of performance score for instructor i at academic year 1998

π_{2i} = the curvature or acceleration in each growth trajectory

e_{it} = the random within-instructor error of prediction for instructor i at time t

The level 2 model

Equation 2.60 characterizes each instructor 's trajectory form academic year 1998 to 2000 by three parameters: π_{0i} , π_{1i} and π_{2i} . The formula of a simple level 2 model (unconditional model) as

$$\pi_{0i} = \beta_{00} + U_{oi} \quad (2.61)$$

$$\pi_{1i} = \beta_{10} + U_{1i} \quad (2.62)$$

$$\pi_{2i} = \beta_{20} \quad (2.63)$$

where π_{0i} = intercept of instructor i, so that given the coding of (AY-1998), π_{oi} is the expected performance score of instructor i at academic year 1998

π_{1i} = the instantaneous rate of change of performance score for instructor i at academic year 1998

π_{2i} = the curvature or acceleration in each growth trajectory

β_{00} = grand mean performance score at academic year 1998

β_{10} = grand mean rate of change of performance score at academic year 1998

β_{20} = grand mean acceleration in performance score

U_{0i} = random effect of instructor i on the performance score at academic year 1998

U_{1i} = random effect of instructor i on the rate of change in performance score

In general, the growth rate at any particular academic year is the first derivative of the growth model evaluated at that academic year. For quadratic growth,

Growth rate at academic year t (G_{it}) = $\pi_{1i} + 2\pi_{2i}$ (AY-1998)$_{it}$ (2.64)

Where π_{1i} = the instantaneous rate of change of performance score for instructor i at academic year 1998

π_{2i} = the curvature or acceleration in each growth trajectory

(AY-1998)$_{it}$ = academic year of instructor i at time t (t=1, 2, 3) minus 1998 so that (AY-1998)$_{it}$ is 0, 1, 2 at academic years 1998, 1999 and 2000

Therefore, the performance score change of instructor i who has non-linear change over a 3 year period were calculated by average growth rate of instructor i at academic years 1998, 1999 and 2000. That is:

$$Q_i = \frac{\sum_{t=1}^{3} G_{it}}{3} \qquad (2.65)$$

where Q_i = the rate of change per academic year in the performance score of instructor i

To conclude, this chapter was composed of 5 sections, which are quality assurance in higher education, tasks and subtasks of the instructors, the development of indicators and weighting, test of linearity and non-linearity and measuring change.

Chapter 3
Materials and Methods

3.1 Introduction

The main objectives of the research are to develop a methodology to measure university instructors' performance scores and also to develop formulae for calculating performance change scores over a 3 year period.

In order to attain the first objective, the major tasks of the instructors were identified and confirmed by senior and experienced staff and also weighted by university's experts. Furthermore, reliability and validity of the performance scores were considered.

In order to attain the second objective, formulae for calculating performance change scores were developed following the consideration of pattern of change and reducing the floor and ceiling effect out off the observed change score. These change scores were validated by score changes obtained from method of Bryk & Raudenbush (1987) and validated by seeking the opinions of heads of department and instructors.

3.2 Selection of Universities for the Case Study

Two Bangkok universities were selected for this study, Thammasat University (**TU**) and Saint John's University (**SU**). They were selected because they introduced Quality Assurance in academic year 1998 and participated in the pilot project for research and development "Indicators, Criteria, and Techniques for Internal and External Assessment in Thai University" run by Prof.Dr.Utumporn Jamornmann on June 2000 - March 2001. Moreover, they are very different kinds of university. TU is a leading public university, inaugurated on June 27, 1934. According to its Website (www.tu.ac.th), it started out as a law school and gradually expanded to offer courses in Social Sciences, Humanities and Sciences (fifteen faculties), with the aim of providing higher education in the fields that are pertinent to social need, and to promote research and Thai culture (TU, 2002). TU also offers graduate programmes leading to master degrees and doctoral degrees. SU, on the other hand, is a private, catholic institution of higher learning, which was granted its charter from Thailand's Ministry of University Affairs on January 5, 1989. According to its Website (www.stjohn.ac.th), it is a part of Saint John's group of nine schools, from kindergarten all the way up to master degrees. SU has five faculties: Business

Administration, Communication Arts, Liberal Arts, Engineering, and Law, and also a Graduate School for master degrees of Business Administration and Communication Arts. The mission of the university is to offer students, in a catholic atmosphere, the opportunity to achieve for themselves a high standard of education and to establish ethics and morals for them to be good Thai citizens (SU, 2002).

There are therefore major differences between the two universities chosen, with SU focussing on teaching and student advice, and TU placing much greater emphasis than SU on research and academic autonomy.

Quality Assurance of both universities is as follows:

a) Thammasat University

Thammasat University (TU) was aware of the significance of Quality Assurance (QA). Therefore QA is included as a main policy since April 1998. Quality Assurance Academic Affairs Department was responsible for TU academic quality under clear objectives and missions of the university on 'TU 4Es' (Excellency, Equity, Ethics and Efficiency). (Quality Assurance Academic Affairs Department, 2002)

TU has done Self Student Report (SSR) and Internal Audit in some faculties, Self Assessment Report (SAR) at faculty and university level, and produced a QA handbook for the university and some faculties such as Science and Technology, Engineering, Medicine, Nursing etc. From Jamornmann's research (2001), the stages of TU 's QA are as follows:

1. Setting a QA committee at university level in October 1998 for steering TU's QA.

2. Identifying principles of QA at university level by TU4Es which reflect the identity and mission of TU.

3. Identifying 9 aspects effecting QA for university and faculty level: 1) philosophy/objectives/implementation, 2) teaching and learning, 3) student development activities, 4) research, 5) academic service to community, 6) preservation of art and culture, 7) administration and management, 8) finance and budgeting, and 9) QA system and mechanisms.

4. Setting up indicators of QA at university level based on 9 aspects.

5. Organizing of Self Study Report (SSR) by faculties for academic year 1999. SSR comprising of 9 aspects was presented in May 2000.

6. Internal audit in 2000 was done for 4 faculties: Medicine on 27-28 April 2000, Nursing on 6-7 July 2000, Engineering on 15-16 August 2000, and Science and Technology on 26-27 September 2000. The aim is to conduct Internal Audit at the faculty level once a year.

7. TU joined the QA pilot project “Indicators, Criteria, and Techniques for Internal and External Assessment in Thai University” run by Prof.Dr. U. Jamornmann. The first meeting was held on 4 July 2000.

8. Jamornmann ’s indicators for Quality Assessment were adapted for use with Self Assessment Report (SAR). University and Faculty ‘s SAR were written and presented to the university and faculty representatives on 15-16 December 2000.

9. SAR at university level was presented to the university council meeting on 22 January 2000 and to the university in the pilot project on 1 March 2000.

b) Saint John’s University

Saint John’s University (SU) has displayed determination in operating a quality policy and as a result the institution has adopted a universal standard. To fulfill this objective every member of personnel, at any level, must be willing to work towards attainment of the universal quality system. Everyone takes part and has responsibility in making a success and sustain the quality system. SU has introduced, developed and is sustaining the quality system. SU has trained and given in service training, seminars for its staff to gain knowledge and understanding of quality systems with the purpose of effective and continuous implementation.

The QA of SU is combined with ISO 9002 (quality system) and Internal QA developed by Prof.Dr.Utumporn’s project as follows:

1) ISO 9002

QA of SU is based on ISO 9002 (quality system) by considering SU ’s teaching and learning process, SU senior staff formed the opinion that ISO 9002 is suitable for SU on the following aspects. (SU, 1999)

1. The system is for quality administration which can be applied to SU ’s process and it is universally certified.

2. The system is for process and document control which is appropriate to SU’s registration and process for teaching/learning materials, text book writing, teaching preparation, evaluation, student data files.

3. ISO 9002 is accepted both inside and outside the country.

4. It is a concrete system which contains data and evidence for reference.

5. There is a monitoring process which pulls out data for the evaluation on the development of a continuous improvement system.

6. It is a system which instructors admire and give their support to.

SU has 7 steps in introducing and operating ISO 9002 QA system as follows:

1. to join the pilot project on Training Lead Consultancy (TLC) on the field of by Bureau of Higher Education Standards, Ministry of University Affairs since April 1998.

2. In September 1998, SU used Quality Manual and trained Internal Auditor (TLC project)

3. In October 1998, SU had Internal Audit.

4. In December 1998, SU began ISO 9002 and finished TLC project.

5. In January 1999, SU trained Internal Auditors, Internal Audit organization and Pre-Assessment Audit at Engineering Institute of Engineering-CBO.

6. In 28 February 1999, SU certified by ISO 9002 from Engineering Institute of Thailand- CBO and was certified by ISO 9002 in 10 May 1999

7. At present time, SU 's operation is still on standard identification by ISO 9000 and quality administration by ISO 9002.

2) Internal Quality Assurance

SU joined Prof.Dr.Jamornmann's Project "Indicators, Criteria, and Techniques for Internal and External Assessment in Thai University" in June 2000. SU has done Self Student Report (SSR), Self Assessment Report (SAR) and Internal Assessment within faculties and offices and SAR at university level. The steps in operation briefly concluded the following. (Jamornmann, 2001)

The first period, SU managed for SSR on the steps as follows:

1. Group of SU staff responsible for quality operation met to identify SSR.

2. The SSR was presented to Quality Representation Committee (QRC) and the Steering Committee (which is composed of Rector and Deans) for approval.

3. A working committee organized a meeting to explain essential points to units / departments of the university.

4. Meetings between the working committee and department representation to set up indicators and criteria for each aspect, starting from a aspect which is simple to implement and not too complicated and finally aspects which are broad and complex thus starting from 6), 5), 4), 3), 1), 8), 9), 2) and 7) respectively. (9 aspects are as follows: 1) philosophy/objectives/implementation, 2) teaching and learning, 3) student development activities, 4) research, 5) academic service to community, 6) preservation of art and culture, 7) administration and management, 8) finance and budgeting, and 9) QA system and mechanisms as same aspects as TU)

5. The working committee considered and reviewed indicators identified in step 4 for the purpose of deciding upon the direction for SSR at department level.

6. Presenting indicators and criteria to senior staff for improvement and after approval from senior staff, presenting them at a meeting to department representatives.

7. In operating SSR at department level, the department assigned people to be responsible for SSR. They might be head of department, instructors or staff of the department.

8. The working committee collected SSR1 from department, and amalgamated the aspects into a single SSR1 for the whole university.

9. The working committee submitted SSR1 to university and analyzed SSR1 in terms of indicators and criteria for each aspect so as to improve and produce SSR2.

The second period is for Self Assessment Report (SAR) and its steps are follows:

1. The working committee called deportment's representatives for a meeting aimed at SAR preparation and changing from SSR to SAR.

2. Faculty/department manage SAR under Dean/Head or Director of working unit responsibility. In this slip each unit assigned an individual responsible for writing SAR.

3. Faculty/department undertakes Self-Assessment within SAR for each aspect.

4. The working committee assessed the SAR from faculties for each aspect.

5. The working committee summarized the assessment.

6. The working committee integrated the reports into the SAR for the university.

7. Presentation of the operation in SU and the universities in the pilot project.

3.3 Data Collecting Groups

Data were collected from 4 groups of academic staff in TU and SU as follows:

Group 1: Task identification and confirmation group was composed of TU and SU senior staff (Academic Vice President, Deputy Deans, and Directors of University's Centers and Officers) and instructors with more than 5 years' experience. They were separated into 2 subgroups:

Subgroup 1.1 Sixteen senior staff and instructors who selected major tasks of the instructor, 7 from TU and 9 from SU. (Please see Appendix B.1)

Subgroup 1.2 Seventeen senior staff and instructors who confirmed the selected subtasks, 8 from TU and 9 from SU. (Please see Appendix B.2)

Group 2: Experts for subtask weighting was composed of 17 senior staff (Academic Vice President, Deputy Dean, and Directors of University's Centers and Officers). They have responsibility for quality assurance and were well aware of instructors' tasks, 9 from TU and 8 from SU. (Please see Appendix B.3)

Group 3: Target Group composed of TU and SU 's instructors in participating faculties who had worked in the faculty since 1998, had never taken more than three months' leave in that time (e.g. personal leave, sick leave, maternity leave, sabbatical, etc.) and had never been a senior administrator such as university president, vice president, dean, deputy dean or director. There were 354 instructors from TU who fit these criteria (Table 3.1) and 60 from SU (Table 3.2), covering a wide range of specializations.

Table 3.1 The Number of TU Instructors, Classifying by Faculty

Faculty	Number of Instructors	Fit Criteria	
		Number	%
1. Law	58	26	44.8
2. Political Science	42	23	54.8
3. Economics	88	38	43.2
4. Social Administration	39	23	59.0
5. Liberal Arts [1]	204	60	29.4
6. Journalism and Mass Communication	45	25	55.6
7. Sociology and Anthropology	32	18	56.3
8. Science and Technology	155	56	36.1
9. Engineering	85	18	21.2
10. Medicine	110	43	39.1
11. Dentistry	43	7	16.3
12. Allied Health Science	30	7	23.3
13. Nursing	24	9	37.5
14. Commerce and Accountancy [2]	114	-	-
15. Language Institution [2]	42	-	-
Total	1,111	354	31.9

[1] Four departments were random from 14 departments

[2] Refused to give data

Table 3.2 The Number of SU Instructors, Classifying by Faculty

Faculty	Number of Instructors	Fit Criteria	
		Number	%
1. Business Administration	68	20	29.4
2. Communication Arts	16	5	31.3
3. Liberal Arts	30	10	33.3
4. Engineering	37	14	37.8
5. Law	5	3	60.0
6. Graduate School	17	8	47.1
Total	173	60	34.7

Group 4: Validation Group was composed of 2 groups: **Group 4.1** contained 13 heads of department, 6 from TU and 7 from SU. **Group 4.2** contained 40 instructors: 20 from each university.

3.4 Data Collection

The researcher collected data from a variety of sources (documents, interviews with university staff, questionnaires etc.), for each of the groups as follows:

3.4.1 Task and Subtask Identification and Confirmation

3.4.1.1 Interviews with Senior Staff

The results of the literature review were used to draft a list of six principal tasks and a list of 64 subtasks that could be used to measure instructors' performance (please see Appendix A for a copy of interview form). Sixteen senior or experienced staff (subgroup 1.1) were asked to comment on the appropriateness of each task and subtask and suggest improvements, deletions and additions.

3.4.1.2 Confirmation of Subtasks

A second group of senior or experienced staff (subgroup 1.2) were used to confirm the appropriateness of the 35 subtasks identified by group 1.1. Each was sent a questionnaire (please see Appendix A for a copy) on which they were asked to rate each of the subtasks on a five point scale to indicate the extent to which they agreed that the subtask was suitable for measuring instructors' performance. The five point scales for considering the appropriate of subtask are as follows:

5 = Strongly agree – Subtask is very appropriate for measuring instructors' performance

4 = Agree – Subtask is appropriate for measuring instructors' performance

3 = Not sure – No idea of appropriateness of subtask for measuring instructors' performance

2 = Disagree – Subtask is not appropriate for measuring instructors' performance

1 = Strongly disagree – Subtask is not at all appropriate for measuring instructors' performance

3.4.2 Weighting of Tasks and Subtasks

Having identified and confirmed the tasks and subtasks that should be used for scoring instructors' performance, the weigh of each task, and of each subtask within each task, needed to be established. Questionnaires were sent to a further group of senior staff (group 2) on which they were asked to (1) identify the weighting of each task on the assumption that the total score of six tasks was 100, and (2) identify the weighting of subtasks in a task on the assumption that the total score of a task was 100 (please see Appendix A for a copy of questionnaire).

3.4.3 Collecting and Scoring Subtasks

Data for the three academic years 1998 (June 1998-May 1999), 1999 (June 1999-May 2000), and 2000 (June 2000-May 2001) were collected from July to October 2001.

Data for 14 of the 35 subtasks could be obtained from university documents such as annual faculty and instructor reports, and course outlines. These data and background details were collected by the researcher (please see table 3.3 for a list of what was collected).

Table 3.3 Content and Sub-content in the Form for Collecting Data from Documents

Content	Sub-content
1. Background	1) Sex
	2) Academic post
	3) Degrees
	4) Teaching experience in higher education
	5) Faculty
	6) Department
2. Teaching	1) Course outlines produced
	2) Teaching load (number of teaching hours per week)
	3) Number of texts used
	4) Number of assessment methods used
	5) Results of evaluation by students
3. Student Advice	1) Number of advisees
4. Research and Academic Publications	1) Number of academic articles published or presented in academic conferences/seminars
	2) Number of research projects contributed to
	3) Number of text or other academic books written
	4) Number of teaching materials written
5. Academic Service to Community	1) Number of invited lecturers (in or out of university)
6. Administration and Academic Self-Improvement	1) Number of permanent committee memberships
	2) Number of temporary committee memberships
	3) Number of times attended academic training/seminars/conferences

The researcher collected data for the 14 “documentary” subtasks. Data for each subtask was classified into five ascending categories scoring 0 to 4 before weighting (please see criteria for scoring in Appendix C).

Data for the other subtasks could only be collected from the instructors themselves. A self-report questionnaire was constructed to facilitate this. Each of the 21 subtasks (Table 3.4) had five options scored 0 to 4 (before weighting). Eight experienced instructors (please see appendix B.4) constructively criticised the questionnaire by improving the options and refining the wording to make it easier to understand. This revised questionnaire was trialled with seven instructors from TU and SU and their feedback was used for further refinements (please see Appendix A for a copy of the questionnaire). Questionnaires were sent to the target group (Group 3). When incomplete questionnaires were returned the researcher telephoned the instructor to obtain the missing data.

Table 3.4 Twenty-one Subtasks in the Self-Report Questionnaire

Task	Subtask
1. Teaching	1) Lesson plans produced
	2) Teaching load (number of teaching hours per week)
	3) Use of appropriate teaching method for students and subjects
	4) Provides activities and environment to stimulate students' learning
	5) Appropriate use of instructional media
	6) Examination administered on time
	7) Grades reported to faculty on time
	8) Use of evaluation results
2. Student Advice	1) Specific times set aside for advising students
	2) Depth of advice offered
	3) Number of advice hours per week
3. Research and Academic Publications	1) Quality of academic work
	2) Status of publications
4. Academic Service to Community	1) Academic service to community on areas of specialization
	2) Academic service to community fits into faculty plan or policy
5. Preservation of Arts and Culture	1) Participant in arts and culture preservation activities
	2) Establishes morals, ethics, and culture into students
	3) Contribution to arts and culture preservation activities / projects for the community
6. Administration and Academic Self-Improvement	1) Attendance at department/faculty/ university meetings
	2) Keeps systematic documentary evidence of work undertaken
	3) Implementation of the knowledge gained from training/seminars/conferences

3.4.4 Validating of the Performance Scores and Performance Change Scores

Thirteen heads of department (group 4.1) were interviewed to verify the validity of the performance scores and performance change scores. They were asked to give their opinion about whether the three yearly performance totals for the instructors in their department correctly reflected instructor performance and their opinion about performance scores change by the methods of both the researcher and Bryk & Raudenbush (1987) in comparing validity of two scores (more, equal, less validity) for his/her department's instructors.

Forty instructors (Group 4.2) were interviewed to verify the validity of their own performance change scores. They were asked to give their opinion about change scores calculated by the researcher's method and that of Bryk & Raudenbush method in comparing validity of two scores (more, equal, less validity).

Table 3.5 Data Collection Classified by Instrument and Data Collecting Group

Data Collecting Group	Instrument	Data Collection
Group 1 Task Identification and Confirmation Group **Subgroup 1.1** Sixteen senior staff and Instructors, 7 from TU and 9 from SU	1. An interview form for selecting major tasks of the instructor	-Researcher interviewed subgroup 1.1 for selecting major tasks of the instructor, getting recommendations for improvement using **Instrument 1**
Subgroup 1.2 Seventeen senior staff and instructors, 8 from TU and 9 from SU	2. An interview form for confirming the subtasks identified	-Researcher interviewed subgroup 1.2 for confirming the appropriateness of subtasks using **Instrument 2**
Group 2 Seventeen experts for identify weighting, 9 from TU and 8 form SU	3. A questionnaire to identify weighting of tasks and subtasks	-Experts identify weighting of 6 tasks for 100 scores and weighting of subtasks in each task for 100 scores using **Instrument 3**

Table 3.5 (Continue)

Data Collecting Group	Instrument	Data Collection
Group 3 Target Group (University instructors), 238 from TU and 52 from SU	4. A form for collecting data from documents	-Researcher collected data of 14 quantitative subtasks from university/faculty documents using **Instrument 4**. For missing data, it was obtained from instructors.
	5. A self-report questionnaire	-Researcher collected data in each subtask which couldn't obtain from documents using **Instrument 5** and getting more data by telephone for incomplete questionnaire
Group 4 Validation Group **Subgroup 4.1** Thirteen heads of department, 6 from TU and 7 from SU	6. An interview form for validating performance scores and performance change scores	-Researcher interviewed heads of department using **Instrument 6**
Subgroup 4.2 Forty instructors, 20 from each university	7. An interview form for validating performance change score	–Researcher interviewed instructors using **Instrument 7**

3.5 Analysis

The researcher analyzed data to answering the research problems that satisfied the objectives.

3.5.1 Based on the research objective 1 "to develop a methodology to measure university instructors' performance scores", the researcher analyzed data as follows:

3.5.1.1 Task and Subtask Identification and Confirmation

a) Task and Subtask identification

Subtasks were identified according to suggestions from 16 senior and experienced staff (subgroup 1.1). Content analysis was performed on the results in order to determine which tasks and subtasks were agreed, which should be combined, which should be removed, what new subtasks should be added and how the definition and wording of the subtasks could be improved. The appropriateness of the six major tasks was confirmed in this way, and 35 subtasks were identified.

b) Confirmation of selected subtasks

Each of the 35 subtasks was rated on a five point scale to indicate appropriateness of subtasks for measuring instructors' performance. The Medians and Quartile Deviations (QD) for each subtask were calculated. Following the rule adopted by Duagmanee (1997), a subtask having a median of not less than 3.50 and with quartile deviation of less than 0.75 was considered to be an appropriate subtask.

3.5.1.2 Weighting of Tasks and Subtasks

Averages for each task and subtask were calculated separately for each university. The weighting for each task at each university was the average thus derived; weightings for each subtask were calculated by multiplying the subtask

average by the relevant task weighting and dividing by 100. Thus the weightings of the subtasks summed to 100.

3.5.1.3 Calculating Performance Scores

The Performance scores for each instructor in academic years 1998, 1999 and 2000 were calculated by multiplying each subtask score by the relevant subtask weighting factor for the instructor's university, and summing the weighted scores that shown in Equation 3.1.

$$Y_i = \sum_{j=1}^{J}\sum_{k=1}^{K} W_{jk} S_{ijk} \qquad (3.1)$$

Where Y_i is the performance score of the i^{th} instructor at the university

$j=1,2,\ldots,6$; j is j^{th} task

$k=1,2,\ldots,K$; k is k^{th} subtask

S_{ijk} is the subtask score of the i^{th} instructor in the k^{th} subtask of the j^{th} task

W_{jk} is the weighting of this university of k^{th} subtask in the j^{th} task

3.5.1.4 Reliability and Validity of Performance Scores

a) Reliability

The following coefficients were calculated separately for each year's scores at each university.

I. Cronbach's alpha. This value, which can range from 0 to 1, indicates the degree to which the subtasks measure the same construct.

II. Pearson correlations between each subtask and the sum of the other subtasks. These values may range from –1 (a perfect, but negative, linear relationship between the subtask and the sum of the others) to 1 (a perfect, positive linear relationship), with 0 indicating no linear relationship. High positive correlations are desirable since they indicate a close match between the subtask and the overall construct being measured.

The subtask-total correlations calculated in stage (II) were used to rank the subtasks. For both universities separately, Spearman's Rho correlations were calculated between the rankings for 1998 and 1999, 1999 and 2000 and 1998 and 2000. These correlations, which may also range from -1 to +1, indicate the degree to which the rankings were the same. High positive values indicate stability in the relationships between subtasks and the total. The three yearly correlations for each subtask at each university were also averaged, and the subtasks ranked by these averages. The Spearman's Rho correlation was calculated between the ranking for SU and TU, so that the stability of the relationship between subtasks across the universities might be investigated.

b) Validity

The opinion of 13 heads of department (subgroup 4.1) about whether the three yearly performance scores for the instructors in their department correctly reflected instructor performance were considered. The percentages of instructors whose scores were confirmed by their Heads as valid or invalid were calculated. The high percentage that scores validated showed concurrent validity.

3.5.2 Based on research objective 2 **"to develop the formulae for calculating performance change scores over a 3 year period", the researcher analyzed data as follows:**

3.5.2.1 Development of formulae for calculating performance change scores

In the last decade, most research has moved beyond the limitation of pre/post (or two waves) measurement. Measuring change with multi-wave data is widely used. The methodology for measurement of change is complex for general people (because of advance methodology ie. advance statistic, measurement theory) and it is difficult to interpret the meaning of change scores. Moreover, it is usually conducted with a large sample size. Therefore, it is necessary to have new formulae that are easy to calculate and translate the meaning of change scores. Furthermore, the formula should calculate change scores for individuals without impact of sample size.

In this research, the academic years during the introduction of quality assurance in TU and SU are 1998, 1999 and 2000. The performance scores over a 3 year period may have linear and non-linear change. Thus, separate formulae for calculating linear and non-linear change scores were developed. The Law of Initial Values (LIV) and Floor and Ceiling Effects with the following concepts were considered when developing the formulae.

"According to the LIV, as scores approach a floor or a ceiling, the same external stimulus should produce smaller changes. Thus the amount of change would be a function of the starting level (i.e., the initial difference from the floor or ceiling). When approaching a ceiling, the higher scores will show a smaller increment, which will result in a negative for r_{dx} (d is change, x is initial value). When approaching a floor, the smaller scores decrease less, which also results in negative r_{dx}." (Jamieson, 1995: 40)

The above concept implies that the performance change score depends on an instructor's performance score especially in the initial year. If an instructor has a high performance score in the initial year; it will be hard for him to increase his score. It shows that the increment of performance score was limited by the ceiling effect.

Besides these, an instructor who had a low performance score in the initial year, his score decreased less than an instructor who had a higher score. It shows that the decrease in performance score was limited by the floor effect. Thus, the performance change score is a function of change of observed scores, initial score, and floor and ceiling effects.

Moreover, there are many patterns of non-linear change (i.e. parabola, quadratic, exponential etc.) in measuring non-linear change. From the concept of organization theory that Sudchari *et al*. (1999) developed from Schewiger *et al*. (1986), that is "Organization survives and grows when it has stability and high adaptability." Any instructor, as a vital part of university, should improve their performance regularly. Thus, the pattern that should have the highest positive change is the pattern that performance score increase regularly; and the pattern that should have the highest negative change is the pattern that performance score decrease regularly. This concept was considered when developing formulae for calculating change score.

From the above concept, formulae for calculating linear or non-linear change score of instructors' performance were developed. There are 2 conditions of the formulae: (1) the data on every subtask should be complete in every academic year and (2) the developed formulae can only be used for calculating change scores on 3 year periods or 3 time points.

3.5.2.2 Test of Linearity and Non-linearity

Performance scores of each instructor was tested for linearity by plotting graph and using analytic geometry theory as follows:

"Let P_1 (X_1, Y_1), P_2 (X_2, Y_2) and P_3 (X_3, Y_3) are the coordinates of academic year (X) and performance scores (Y) in 3 academic years and let slope of P_1P_2 is m_1 and slope of P_1P_3 is m_2. The points P_1, P_2 and P_3 will be lie on the line (linear) if $m_1=m_2$, where $m_1 = \frac{y_1 - y_2}{x_1 - x_2}$ and $m_2 = \frac{y_1 - y_3}{x_1 - x_3}$" (Fuller, 1967: 28-29) (please see details in section 2.4, chapter 2)

3.5.2.3 Validity of Change Scores obtained from the researcher's formula

The change scores using the formula developed by the researcher were validated by 2 methods:

a) Correlating With Change Scores Obtained from Bryk & Raudenbush (1987)

Performance change scores of each instructor were calculated by 2 methods: Researcher's method and Bryk & Raudenbush (1987) method. The method of Bryk

& Raudenbush (1987) for measuring change was used as criterion. Pearson 's product moment correlations (r_{XY}) were used to verify the concurrent validity (Equation 3.2).

$$r_{XY} = \frac{\sum_{i=1}^{N}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{N}(X_i - \overline{X})^2 \sum_{i=1}^{N}(Y_i - \overline{Y})^2}} \qquad (3.2) \quad \text{(Neter, 1992 : 558)}$$

Where X_i = The change score of instructor i obtained form the researcher's formula

Y_i = The change score of instructor i obtained form the Bryk & Raudenbush's formula

N = Number of instructors

The values of correlation can range from –1 (a perfect, but negative, linear relationship) to 1 (a perfect, positive linear relationship), with 0 indicating no linear relationship. The high positive correlation (obtained from Equation 3.2) indicates the change scores obtained from research's formula have concurrent validity.

b) Seeking Opinion of Heads of Department and Instructors

The other validity was obtained by seeking the opinions of heads of department and instructors by comparing the score changes of individual instructors obtained by formula developed by the researcher with those of Bryk & Raudenbush (1987). The percentage of opinions was considered. If high percentage of heads of department and instructors' opinions agreed that change scores obtained by the researcher's formula were more valid, they appear to have greater validity than those obtained by Bryk & Raudenbush (1987).

Chapter 4
Results

4.1 Introduction

Thammasat University (TU) and Saint John's University (SU) were the case study university in this research. The researcher received very good cooperation of TU and SU staff for collecting data from their faculty and university documents and themselves. Almost all the senior staff at both universities who were asked about appropriateness and weighting of subtasks and the validity of the scores returned their questionnaires. The self-report questionnaires sent to instructors had a lower, but still high, response rate (67% at TU and 87% at SU) (please see table 4.1 and 4.2). The higher response rate at SU was probably because the Academic Vice President asked instructors to complete the questionnaire and return them to him, whereas at TU, with its tradition of faculty and instructor autonomy, the researcher wrote to instructors herself. This illustrates the usefulness of obtaining active support from senior staff. Nevertheless, the 67% response rate for TU is still high, probably because the researcher provided stamped, addressed return envelopes and send up to two reminders to instructors who did not reply straight away.

The problems of data collection were that some faculties didn't have systematic documentary evidence of their instructors' work and some faculties in TU refused to give data.

Table 4.1 Percentage of Data Received From TU 's Instructors Classifying by Faculty

Faculty	Number of Instructors Fit Criteria	Data Received	
		Number	%
1. Law	26	16	62
2. Political Science	23	12	52
3. Economics	38	20	53
4. Social Administration	23	19	83
5. Liberal Arts	60	45	75
6. Journalism and Mass Communication	25	12	48
7. Sociology and Anthropology	18	8	44
8. Science and Technology	56	39	70
9. Engineering	18	15	83
10. Medicine	43	29	67
11. Dentistry	7	7	100
12. Allied Health Science	7	7	100
13. Nursing	9	9	100
Total	354	238	67

Table 4.2 Percentage of Data Received From SU 's Instructors Classifying by Faculty

Faculty	Number of Instructors Fit Criteria	Data Received	
		Number	%
1. Business Administration	20	17	85
2. Communication Arts	5	4	80
3. Liberal Arts	10	9	90
4. Engineering	14	13	93
5. Law	3	3	100
6. Graduated School	8	6	75
Total	60	52	87

4.2 Results of Measuring University Instructor's Performance Scores

compose of 4 past as follows:

4.2.1 Tasks and subtasks of university instructors

4.2.2 Weighting of tasks and subtasks

4.2.3 Instructor's performance scores

4.2.4 Reliability and validity of performance scores

4.2.1 Tasks and Subtasks of University Instructors

The results of the literature review were used to draft a list of six principal tasks consisting of teaching, student advice, research and academic publication, academic service to community, preservation of arts and culture, and administration and academic self-improvement. 64 subtasks were identified, 42 of which were quantitative and 22 of which were qualitative. See table 2.3 (section 2.2, p.41-42) for a list of the subtasks identified. Sixteen senior or experienced staff (subgroup 1.1), were asked to comment on the appropriateness of each task and subtask and suggest improvements, deletions and additions. Content analysis was performed on the results in order to determine which tasks and subtasks were agreed, which should be combined, which should be removed, what new subtasks should be added and how the definition and wording of the subtasks could be improved. The appropriateness of the six major tasks was confirmed in this way, whilst the 64 initial subtasks were reduced to 35, 21 of which were quantitative and 14 of which were qualitative (see the first column of Table 4.3).

A second group of senior or experienced staff (subgroup 1.2) was used to confirm the appropriateness of the subtasks identified. Each was sent a questionnaire on which they were asked to rate each of the 35 subtasks on a five point scale to indicate the extent to which they agreed that the subtask was suitable for measuring instructors' performance. The medians and quartile deviations (QD) of the results for each subtask are shown in the "Appropriateness of subtasks" columns of Table 4.3. Following the rule adopted by Duagmanee (1997), a subtask having a median of not less than 3.50 and with quartile deviation of less than 0.75 was

considered to be an appropriate subtask. The medians of subtasks for TU ranged from 3.50 - 5.00 and the QDs from 0.00 - 0.63; for SU the medians ranged from 4.00 - 5.00 and QDs from 0.00 - 5.00. Therefore all the subtasks were confirmed as appropriate at both universities.

Table 4.3 Tasks and Subtasks of University Instructors

Tasks and Subtasks	Subtask Type	Appropriateness of subtask [1]			
		TU		SU	
		Median	QD	Median	QD
Task 1: Teaching					
1.1 Course outlines produced	Quantity	5.00	0.00	5.00	0.50
1.2 Lesson plans produced	Quantity	5.00	0.13	5.00	0.50
1.3 Teaching load (number of teaching hours per week)	Quantity	4.50	0.50	4.00	0.50
1.4 Use of appropriate teaching method for students and subjects	Quality	4.50	0.63	5.00	0.50
1.5 Provides activities and environment that stimulate students' learning	Quality	4.50	0.63	5.00	0.50
1.6 Appropriate use of instructional media	Quality	4.00	0.63	4.00	0.50
1.7 Number of texts used	Quantity	3.50	0.63	4.00	0.50
1.8 Number of assessment methods used	Quantity	3.50	0.50	4.00	0.50
1.9 Examination administered on time	Quantity	3.50	0.63	4.00	0.50
1.10 Grades reported to faculty on time	Quantity	4.50	0.50	5.00	0.50
1.11 Results of evaluation by students	Quantity	4.00	0.63	5.00	0.00
1.12 Use of evaluation results	Quality	4.50	0.50	5.00	0.00
Task 2: Student Advice					
2.1 Specific times set aside for advising students	Quality	4.00	0.13	4.00	0.50
2.2 Depth of advice offered	Quality	4.00	0.25	5.00	0.50
2.3 Number of advice hours per week	Quantity	4.00	0.63	4.00	0.50
2.4 Number of advisees	Quantity	4.00	0.63	4.00	0.50
Task 3: Research and Academic Publications					
3.1 Number of academic articles published or presented in academic conferences/seminars	Quantity	4.50	0.50	4.00	0.50
3.2 Number of research projects contributed to	Quantity	4.00	0.00	4.00	0.50
3.3 Number of text or other academic books written	Quantity	4.00	0.50	4.00	0.50
3.4 Number of teaching materials written	Quantity	4.00	0.13	4.00	0.00
3.5 Quality of academic work	Quality	4.00	0.25	5.00	0.50
3.6 Status of publications	Quality	4.00	0.00	5.00	0.50
Task 4: Academic Service to Community					
4.1 Number of invited lecturers (in or out university)	Quantity	4.00	0.50	4.00	0.50
4.2 Number of service hours per week	Quantity	4.00	0.50	4.00	0.50
4.3 Academic service to community on areas of specialization	Quality	4.00	0.63	4.00	0.50
4.4 Academic service to community fits into faculty plan or policy	Quality	4.00	0.13	4.00	0.50
Task 5: Preservation of Arts and Culture					
5.1 Participant in arts and culture preservation activities	Quantity	4.00	0.63	4.00	0.50
5.2 Establishes morals, ethics, and culture into students	Quality	4.00	0.50	5.00	0.50
5.3 Contribution to arts and culture preservation activities / projects for the community	Quality	3.50	0.50	4.00	0.50
Task 6: Administration and Academic Self-Improvement					
6.1 Number of permanent committee memberships	Quantity	3.50	0.63	4.00	0.50
6.2 Number of temporary committee memberships	Quantity	4.00	0.50	4.00	0.50
6.3 Attendance at department/faculty/ university meetings	Quantity	3.50	0.50	4.00	0.50
6.4 Keeps systematic documentary evidence of work undertaken	Quality	3.50	0.50	4.00	0.50
6.5 Number of times attended academic training/seminars/conferences	Quantity	4.00	0.63	4.00	0.50
6.6 Implementation of the knowledge gained from training/seminars/conferences	Quality	4.00	0.25	5.00	0.00

[1] a subtask having a median of not less than 3.50 and with quartile deviation of less than 0.75 was considered to be an appropriate subtask.

4.2.2 Weighting of Tasks and Subtasks

Having identified and confirmed the tasks and subtasks that should be used for scoring instructors' performance, the weigh of each task, and of each subtask within each task, needed to be established. Questionnaires were sent to a further group of senior staff (group 2) to identify the weighting of tasks and subtasks. The weightings for each task and subtask at each university were calculated according to the method described in section 3.5.1.2. The results are shown in Table 4.4.

Table 4.4. The Weighting of Tasks and Subtasks

Tasks and Subtasks	Weighting (Total=100)	
	TU	SU
Task 1: Teaching	**38**	**38**
1.1 Course outlines produced	3.0	2.3
1.2 Lesson plans produced	3.5	3.3
1.3 Teaching load (number of teaching hours per week)	3.6	3.1
1.4 Use of appropriate teaching method for students and subjects	4.4	4.3
1.5 Provides activities and environment that stimulate students' learning	4.1	4.2
1.6 Appropriate use of instructional media	3.8	4.1
1.7 Number of texts used	3.4	2.8
1.8 Number of assessment methods used	2.1	2.5
1.9 Examination administered on time	2.2	2.5
1.10 Grades reported to faculty on time	2.6	2.5
1.11 Results of evaluation by students	2.5	2.6
1.12 Use of evaluation results	2.8	3.8
Task 2: Student Advice	**10**	**16**
2.1 Specific times set aside for advising students	2.1	3.1
2.2 Depth of advice offered	3.8	5.7
2.3 Number of advice hours per week	2.1	3.6
2.4 Number of advisees	2.0	3.6
Task 3: Research and Academic Publications	**23**	**15**
3.1 Number of academic articles published or presented in academic conferences/seminars	4.0	2.2
3.2 Number of research projects contributed to	4.0	2.9
3.3 Number of text or other academic books written	5.1	2.9
3.4 Number of teaching materials written	3.4	2.6
3.5 Quality of academic work	3.3	2.2
3.6 Status of publications	3.2	2.2
Task 4: Academic Service to Community	**12**	**9**
4.1 Number of invited lecturers (in or out university)	2.7	2.0
4.2 Number of service hours per week	2.7	1.8
4.3 Academic service to community on areas of specialization	3.6	2.5
4.4 Academic service to community fits into faculty plan or policy	3.0	2.7
Task 5: Preservation of Arts and Culture	**6**	**9**
5.1 Participant in arts and culture preservation activities	1.4	2.3
5.2 Establishes morals, ethics, and culture into students	2.9	4.5
5.3 Contribution to arts and culture preservation activities/projects for the community	1.7	2.2
Task 6: Administration and Academic Self-Improvement	**11**	**13**
6.1 Number of permanent committee memberships	2.0	1.6
6.2 Number of temporary committee memberships	1.5	1.7
6.3 Attendance at department/faculty/ university meetings	1.8	2.1
6.4 Keeps systematic documentary evidence of work undertaken	1.1	1.6
6.5 Number of times attended academic training/seminars/conferences	2.0	2.2
6.6 Implementation of the knowledge gained from training/seminars/conferences	2.6	3.8

Table 4.4 shows both universities considered teaching to have highest weight. TU give eight points more weight to research and academic publications than SU, whilst Su give six more points than TU to student advice.

The weighting of subtasks (total 100) for TU range from 1.1 (subtask 6.4, keeps systematic documentary evidence of work undertaken) to 5.1 (subtask 3.3,number of text or other academic books written), whilst SU range from 1.6 (subtask 6.4, keeps systematic documentary evidence of work undertaken and subtask 6.1, number of permanent committee memberships) to 5.7 (subtask 2.2, Depth of advice offered).

4.2.3 Instructors' Performance Scores

Questionnaires were sent to the target group (group 3). Questionnaires were returned by 238 TU instructors (a 67% response rate) and 52 SU instructors (87%). Background information is shown in Table 4.5.

From Table 4.5, the majority of the 238 TU instructors who gave their data were female (60%). The most common highest qualification was a masters' degree (60%), and the most common position was lecturer (41%). The faculty contributing the largest number of instructors was Liberal Arts (19%). On average, the instructors had 14 years' experience.

The majority of the 52 SU instructors who gave their data were male (63%). 94% were lecturers, and 75% had masters' degrees. 33% worked in the faculty of Business and Administration. On average, they had 8 years' experience.

Table 4.5 Background of TU and SU instructors

Background	TU		SU	
	Number	%	Number	%
1. Sex				
- Male	96	40.34	33	63.46
- Female	142	59.66	19	36.54
Total	**238**	**100.00**	**52**	**100.00**
2. Degree				
- Bachelor	11	4.62	9	17.31
- Master	142	59.66	39	75.00
- Doctor	85	35.71	4	7.69
Total	**238**	**100.00**	**52**	**100.00**
3. Academic Post				
- Lecturer	97	40.76	49	94.23
- Assistant Professor	77	32.35	3	5.77
- Associate Professor	60	25.21	0	0.00
- Professor	4	1.68	0	0.00
Total	**238**	**100.00**	**52**	**100.00**
4. Teaching Experience in Higher Education				
Less than 6 years	56	23.53	17	32.70
6 – 10 years	58	24.37	31	59.62
11 – 15 years	29	12.18	1	1.92
16 – 20 years	31	13.03	1	1.92
21 – 25 years	24	10.08	0	0.00
26 – 30 years	28	11.76	0	0.00
More than 30 years	12	5.04	2	3.84
Total	**238**	**100.00**	**52**	**100.00**
	$\overline{X}$ = 14, SD = 9.57		$\overline{X}$ = 8, SD = 6.70	

Table 4.5 (Continue)

Background	TU		SU	
	Number	%	Number	%
5. Faculty				
-Law	16	6.72	3	5.77
-Political Science	12	5.04	-	-
-Economics	20	8.40	-	-
-Social Administration	19	7.98	-	-
-Liberal Arts	45	18.91	9	17.31
-Journalism and Mass Communication	12	5.04	-	-
-Sociology and Anthropology	8	3.36	-	-
-Science and Technology	39	16.39	-	-
-Engineering	15	6.30	13	25.00
-Medicine	29	12.18	-	-
-Dentistry	7	2.94	-	-
-Allied Health Science	7	2.94	-	-
-Nursing	9	3.78	-	-
-Business Administration	-	-	17	32.69
-Communication Arts	-	-	4	7.69
-Graduate School	-	-	6	11.54
Total	**238**	**100.00**	**52**	**100.00**

A performance score was calculated for each instructor by multiplying each subtask score (ranging from 0 to 4) by the relevant subtask weighting factor for the instructor's university, and summing the weighted scores (see section 3.5.1.3). Task and subtask weighting factors and maximum weighted scores are shown in columns two and three of Tables 4.6 and 4.7. Total performance scores may range from 0 to 400. Means and standard deviations for each of the three academic years of performance scores are shown in Tables 4.6 and 4.7.

Table 4.6 Mean and Standard Deviation of TU Instructors' Performance Scores (n=238), Classifying by Academic Year (AY)

Tasks and Subtasks	Weight	Instructors' performance scores							
		Max	AY 1998		AY 1999		AY 2000	
			Mean	SD	Mean	SD	Mean	SD
Task 1: Teaching	**38**	**152**	**95.80**	**21.18**	**102.32**	**18.53**	**106.04**	**17.84**
1.1 Course outlines produced	3.0	12	10.50	3.05	10.92	2.30	11.18	2.07
1.2 Lesson plans produced	3.5	14	9.25	4.67	9.88	4.23	10.28	4.17
1.3 Teaching load (number of teaching hours per week)	3.6	14.4	5.38	3.32	6.55	3.62	6.78	3.73
1.4 Use of appropriate teaching method for students and subjects	4.4	17.6	9.30	5.35	10.82	5.01	12.24	5.17
1.5 Provides activities and environment that stimulate students' learning	4.1	16.4	8.56	4.71	9.25	4.39	9.47	4.40
1.6 Appropriate use of instructional media	3.8	15.2	10.59	4.04	11.13	3.63	11.37	3.54
1.7 Number of texts used	3.4	13.6	6.61	5.06	6.91	5.03	7.13	5.07
1.8 Number of assessment methods used	2.1	8.4	4.50	2.96	4.67	2.93	4.77	2.87
1.9 Examination administered on time	2.2	8.8	7.75	1.77	7.86	1.54	7.92	1.52
1.10 Grades reported to faculty on time	2.6	10.4	8.98	2.14	9.14	1.94	9.22	1.98
1.11 Results of evaluation by students	2.5	10	7.57	2.99	8.07	2.71	8.34	2.65
1.12 Use of evaluation results	2.8	11.2	6.80	2.26	7.13	2.13	7.34	2.18
Task 2: Student Advice	**10**	**40**	**23.17**	**7.85**	**24.29**	**7.40**	**24.82**	**7.38**
2.1 Specific times set aside for advising students	2.1	8.4	4.87	3.40	4.99	3.37	5.16	3.36
2.2 Depth of advice offered	3.8	15.2	10.14	3.68	10.39	3.53	10.52	3.49
2.3 Number of advice hours per week	2.1	8.4	4.39	2.38	4.50	2.41	4.54	2.45
2.4 Number of advisees	2.0	8	3.76	2.99	4.40	3.08	4.60	3.12

Table 4.6 (Continue)

Tasks and Subtasks	Weight	Instructors' performance scores							
		Max	AY 1998		AY 1999		AY 2000	
			Mean	SD	Mean	SD	Mean	SD
Task 3: Research and Academic Publications	**23**	**92**	**20.62**	**13.12**	**22.64**	**13.64**	**25.28**	**14.20**
3.1 Number of academic articles published or presented in academic conferences/seminars	4.0	16	3.11	4.18	3.43	4.60	3.93	4.83
3.2 Number of research projects contributed to	4.0	16	1.51	3.05	1.85	3.26	1.93	3.13
3.3 Number of text or other academic books written	5.1	20.4	1.76	3.34	1.80	3.63	2.61	4.18
3.4 Number of teaching materials written	3.4	13.6	2.73	3.71	3.04	3.85	3.40	4.05
3.5 Quality of academic work	3.3	13.2	5.88	3.31	6.43	3.18	7.00	3.16
3.6 Status of publications	3.2	12.8	5.63	3.68	6.09	3.69	6.40	3.67
Task 4: Academic Service to Community	**12**	**48**	**23.43**	**11.73**	**24.48**	**11.27**	**25.07**	**11.20**
4.1 Number of invited lecturers (in or out university)	2.7	10.8	3.06	3.44	3.27	3.60	3.58	3.63
4.2 Number of service hours per week	2.7	10.8	3.46	3.57	3.65	3.56	3.66	3.54
4.3 Academic service to community on areas of specialization	3.6	14.4	9.42	4.76	9.76	4.48	9.95	4.38
4.4 Academic service to community fits into faculty plan or policy	3.0	12	7.49	3.87	7.80	3.68	7.87	3.71
Task 5: Preservation of Arts and Culture	**6**	**24**	**10.17**	**5.27**	**10.87**	**5.17**	**11.09**	**5.29**
5.1 Participant in arts and culture preservation activities	1.4	5.6	2.10	1.45	2.12	1.42	2.14	1.41
5.2 Establishes morals, ethics, and culture into students	2.9	11.6	6.37	3.30	6.88	3.18	7.07	3.25
5.3 Contribution to arts and culture preservation activities / projects for the community	1.7	6.8	1.70	1.85	1.86	1.88	1.89	1.94
Task 6: Administration and Academic Self-Improvement	**11**	**44**	**23.21**	**6.52**	**25.54**	**6.25**	**26.22**	**6.64**
6.1 Number of permanent committee memberships	2.0	8	2.82	2.55	3.56	2.49	3.82	2.57
6.2 Number of temporary committee memberships	1.5	6	2.38	2.05	2.72	2.04	2.85	2.12
6.3 Attendance at department/faculty/ university meetings	1.8	7.2	4.75	1.71	4.73	1.70	4.73	1.71
6.4 Keeps systematic documentary evidence of work undertaken	1.1	4.4	2.60	1.01	2.75	0.90	2.82	0.90
6.5 Number of times attended academic training/seminars/conferences	2.0	8	3.50	2.75	4.03	2.72	4.18	2.80
6.6 Implementation of the knowledge gained from training/seminars/conferences	2.6	10.4	7.16	2.51	7.75	2.02	7.81	2.09
Total	**100**	**400**	**196.40**	**42.61**	**210.15**	**38.11**	**218.53**	**37.92**

[1] Weighting of tasks and subtasks obtained by TU experts
[2] Maximum of TU instructors' performance score in each task and subtask

Table 4.6 shows the mean performance score of TU instructors in academic years 1998, 1999 and 2000 were 196, 210 and 219 respectively.

There was a gradual increase in the mean score at TU of each subtask across the three years, apart from subtask 6.3, attendance at department / faculty/ university meetings, which essentially stayed the same. If the means for the final year are expressed as a percentage of the maximum score available, they range from 12% (subtask 3.2, number of research projects contributed to) to 93% (subtask 1.1, course outlines produced), with a mean of 53% and standard deviation of 21%.

Table 4.7 Mean and Standard Deviation of SU Instructors' Performance Scores (n=52), Classifying by Academic Year (AY)

Tasks and Subtasks	Weight	Instructors' performance scores							
		Max	AY 1998		AY 1999		AY 2000	
			Mean	SD	Mean	SD	Mean	SD
Task 1: Teaching	**38**	**152**	**82.77**	**24.00**	**96.38**	**20.32**	**105.87**	**19.19**
1.1 Course outlines produced	2.3	9.2	6.50	3.01	7.25	2.51	7.70	2.23
1.2 Lesson plans produced	3.3	13.2	6.79	4.73	8.63	4.29	9.52	4.32
1.3 Teaching load (number of teaching hours per week)	3.1	12.4	5.19	3.92	5.84	4.24	6.44	4.02
1.4 Use of appropriate teaching method for students and subjects	4.3	17.2	6.78	4.99	10.34	4.28	13.48	4.00
1.5 Provides activities and environment that stimulate students' learning	4.2	16.8	7.59	4.40	9.69	3.86	10.66	4.03
1.6 Appropriate use of instructional media	4.1	16.4	9.70	4.30	11.20	3.46	11.98	3.33
1.7 Number of texts used	2.8	11.2	2.96	3.16	3.34	3.33	4.04	3.65
1.8 Number of assessment methods used	2.5	10	4.47	2.59	4.42	2.87	5.00	2.89
1.9 Examination administered on time	2.5	10	7.74	2.72	8.13	2.37	8.46	2.28
1.10 Grades reported to faculty on time	2.5	10	8.27	2.50	8.61	2.12	8.70	2.13
1.11 Results of evaluation by students	2.6	10.4	7.50	2.90	8.70	2.12	9.15	2.09
1.12 Use of evaluation results	3.8	15.2	9.28	2.65	10.23	2.20	10.74	2.22
Task 2: Student Advice	**16**	**64**	**46.32**	**13.04**	**49.43**	**12.19**	**49.85**	**11.78**
2.1 Specific times set aside for advising students	3.1	12.4	8.35	4.10	9.54	3.46	9.78	3.38
2.2 Depth of advice offered	5.7	22.8	16.44	5.13	18.09	4.88	18.20	4.93
2.3 Number of advice hours per week	3.6	14.4	10.80	3.70	11.01	3.73	11.22	3.32
2.4 Number of advisees	3.6	14.4	10.73	5.68	10.80	5.57	10.66	5.52
Task 3: Research and Academic Publications	**15**	**60**	**8.24**	**7.09**	**10.13**	**8.05**	**12.08**	**8.43**
3.1 Number of academic articles published or presented in academic conferences/seminars	2.2	8.8	0.97	1.92	1.10	1.87	1.35	2.14
3.2 Number of research projects contributed to	2.9	11.6	0.33	1.37	0.56	1.82	0.67	1.87
3.3 Number of text or other academic books written	2.9	11.6	0.45	1.20	0.50	1.60	0.84	1.75
3.4 Number of teaching materials written	2.6	10.4	1.45	2.21	1.75	2.45	2.20	2.88
3.5 Quality of academic work	2.2	8.8	2.67	2.15	3.34	1.92	3.77	1.92
3.6 Status of publications	2.2	8.8	2.37	2.38	2.88	2.20	3.26	2.28
Task 4: Academic Service to Community	**9**	**36**	**12.45**	**7.86**	**14.57**	**8.09**	**15.62**	**7.74**
4.1 Number of invited lecturers (in or out university)	2.0	8	1.23	2.10	1.38	2.08	1.58	2.11
4.2 Number of service hours per week	1.8	7.2	1.56	1.92	2.22	2.27	2.32	2.35
4.3 Academic service to community on areas of specialization	2.5	10	4.47	3.11	5.05	2.91	5.34	2.76
4.4 Academic service to community fits into faculty plan or policy	2.7	10.8	5.19	3.16	5.92	3.03	6.39	2.73
Task 5: Preservation of Arts and Culture	**9**	**36**	**16.75**	**6.80**	**18.61**	**7.15**	**19.34**	**7.86**
5.1 Participant in arts and culture preservation activities	2.3	9.2	4.78	2.18	5.04	2.09	5.13	2.17
5.2 Establishes morals, ethics, and culture into students	4.5	18	9.69	4.31	10.73	4.38	11.25	4.76
5.3 Contribution to arts and culture preservation activities / projects for the community	2.2	8.8	2.28	2.26	2.83	2.36	2.96	2.61
Task 6: Administration and Academic Self-Improvement	**13**	**52**	**21.96**	**8.59**	**25.70**	**8.64**	**28.46**	**9.79**
6.1 Number of permanent committee memberships	1.6	6.4	1.66	1.55	2.58	2.08	3.29	2.34
6.2 Number of temporary committee memberships	1.7	6.8	2.09	2.01	2.65	1.89	2.78	2.05
6.3 Attendance at department/faculty/ university meetings	2.1	8.4	4.56	2.55	4.89	2.41	5.01	2.50
6.4 Keeps systematic documentary evidence of work undertaken	1.6	6.4	3.72	1.67	4.15	1.46	4.43	1.51

Table 4.7 (Continue)

Tasks and Subtasks	Weight	Instructors' performance scores							
		Max	AY 1998		AY 1999		AY 2000	
			Mean	SD	Mean	SD	Mean	SD
6.5 Number of times attended academic training/seminars/conferences	2.2	8.8	1.44	2.34	1.27	2.06	2.28	2.82
6.6 Implementation of the knowledge gained from training/seminars/conferences	3.8	15.2	8.48	4.10	10.16	3.82	10.67	3.84
Total	**100**	**400**	**188.49**	**47.89**	**214.81**	**41.66**	**231.23**	**44.59**

Note [1] Weighting of tasks and subtasks obtained by SU experts
[2] Maximum of SU instructors' performance score in each task and subtask

Table 4.7 shows the mean performance score of SU instructors in academic years 1998, 1999 and 2000 were 188, 215 and 231 respectively.

Almost all the subtasks showed a gradual increase in their mean values at SU, the only exceptions being subtask 1.8 (number of assessment methods used) and 6.5 (number of times attended academic training/seminars/conferences), which both declined slightly between the first two years, and subtask 2.4 (number of advisees), which declined a little between years two and three. Final year mean scores expressed as percentages of the maximum available range from 6% (subtask 3.2, number of research projects contributed to – the same subtask was lowest for TU) to 88% (subtask 1.11, results of evaluation by students), with a mean of 55% and standard deviation of 24%.

4.2.4 Reliability and Validity of Performance Scores

Reliability

Table 4.8 shows Cronbach's alpha for each year and university. The values, which range between 0.76 and 0.81 at TU and 0.84 and 0.87 at SU, are very slightly higher at SU.

Table 4.8 Reliability of the scores

	TU (n=238)			**SU** (n=52)		
	1998	1999	2000	1998	1999	2000
Cronbach's alpha	.81	.77	.76	.87	.84	.86

Table 4.9 shows the correlation between each subtask and the total of the other subtasks. These figures represent the degree to which there is a linear relationship between the subtask and the sum of the other subtasks. Values may range from –1 (a perfect, but negative, linear relationship) to 1 (a perfect, positive linear relationship), with 0 indicating no linear relationship.

The most striking feature about Table 4.9 is the similarity of the correlations across the years. If the subtasks are ranked by the correlations, then the Spearman's Rho correlation between the rankings for the different years in TU range from 0.88 (between 1998 and 2000) to 0.96 (between 1999 and 2000); for SU they range from 0.81 (between 1998 and 1999) to 0.90 (between 1999 and 2000).

If the subtasks are ranked by the three-yearly average of their correlations at each university, Spearman's Rho between TU and SU is 0.42.

Turning briefly to the actual figures in Table 4.9, the mean correlation of all the subtasks each year was around 0.3 for TU, and 0.4 for SU. Almost all correlations were positive, and none were strongly negative. A couple of subtasks had correlations that were consistently very low at both universities, indicating that there was little relationship between them and instructor performance scores. These were subtasks 2.4 (number of advisees, i.e. the number of students advised by the instructor) and 3.2 (number of research projects contributed to). 3.2's correlation for SU actually started slightly negative and moved closer to 0 each year (the average across the 3 years for this subtask was -0.09 at SU, compared with +0.09 at TU). All the subtasks were retained, however, because they had all been endorsed by senior staff from each university (i.e. they were needed for validity), none of the correlations were strongly negative, most were above 0 and Cronbach's alpha was acceptable.

Table 4.9 Correlation Between Subtask and The Total of The Other Subtasks

Tasks and Subtasks	**TU** (n=238)			**SU** (n=52)		
	1998	1999	2000	1998	1999	2000
Task 1: Teaching						
1.1 Course outlines produced	0.38	0.28	0.23	0.45	0.36	0.32
1.2 Lesson plans produced	0.36	0.32	0.24	0.36	0.30	0.31
1.3 Teaching load (number of teaching hours per week)	-0.05	-0.10	-0.04	0.27	0.43	0.30
1.4 Use of appropriate teaching method for students and subjects	0.54	0.53	0.51	0.48	0.41	0.47
1.5 Provides activities and environment that stimulate students' learning	0.52	0.48	0.48	0.45	0.44	0.42
1.6 Appropriate use of instructional media	0.24	0.19	0.23	0.48	0.43	0.45
1.7 Number of texts used	0.25	0.21	0.17	0.51	0.31	0.26
1.8 Number of assessment methods used	0.09	0.11	0.12	0.28	0.05	0.23
1.9 Examination administered on time	0.20	0.11	0.05	0.59	0.58	0.56
1.10 Grades reported to faculty on time	0.19	0.07	0.02	0.63	0.61	0.58
1.11 Results of evaluation by students	0.34	0.25	0.26	0.55	0.53	0.57
1.12 Use of evaluation results	0.45	0.45	0.45	0.50	0.45	0.40
Task 2: Student Advice						
2.1 Specific times set aside for advising students	0.41	0.35	0.31	0.45	0.35	0.42
2.2 Depth of advice offered	0.40	0.36	0.37	0.53	0.53	0.51
2.3 Number of advice hours per week	0.34	0.34	0.30	0.47	0.29	0.40
2.4 Number of advisees	-0.01	0.00	-0.03	0.06	-0.01	0.13
Task 3: Research and Academic Publications						
3.1 Number of academic articles published or presented in academic conferences/seminars	0.25	0.27	0.27	0.35	0.28	0.25
3.2 Number of research projects contributed to	0.07	0.10	0.10	-0.16	-0.08	-0.03
3.3 Number of text or other academic books written	0.20	0.12	0.12	0.08	0.12	0.13
3.4 Number of teaching materials written	0.28	0.19	0.23	0.00	0.13	0.11
3.5 Quality of academic work	0.42	0.41	0.41	0.56	0.46	0.54
3.6 Status of publications	0.34	0.34	0.33	0.36	0.38	0.49
Task 4: Academic Service to Community						
4.1 Number of invited lectures (in or out of university)	0.39	0.36	0.31	0.27	0.36	0.37
4.2 Number of service hours per week	0.20	0.21	0.22	0.31	0.37	0.41
4.3 Academic service to community on areas of specialization	0.46	0.34	0.32	0.53	0.59	0.64
4.4 Academic service to community fits into faculty plan or policy	0.51	0.43	0.38	0.48	0.57	0.64

Table 4.9 (Continue)

Tasks and Subtasks	**TU** (n=238)			**SU** (n=52)		
	1998	1999	2000	1998	1999	2000
Task 5: Preservation of Arts and Culture						
5.1 Participant in arts and culture preservation activities	0.28	0.32	0.34	0.61	0.47	0.60
5.2 Establishes morals, ethics, and culture into students	0.47	0.40	0.40	0.39	0.36	0.40
5.3 Contribution to arts and culture preservation activities / projects for the community	0.38	0.38	0.40	0.34	0.29	0.39
Task 6: Administration and Academic Self-Improvement						
6.1 Number of permanent committee memberships	0.28	0.31	0.38	0.44	0.32	0.41
6.2 Number of temporary committee memberships	0.22	0.22	0.29	0.37	0.15	0.23
6.3 Attendance at department/faculty/ university meetings	0.20	0.16	0.15	0.47	0.45	0.50
6.4 Keeps systematic documentary evidence of work undertaken	0.32	0.31	0.33	0.44	0.38	0.33
6.5 Number of times attended academic training/seminars/conferences	0.21	0.17	0.17	0.27	0.21	0.21
6.6 Implementation of the knowledge gained from training/seminars/conferences	0.40	0.38	0.33	0.37	0.39	0.39
Mean	**0.30**	**0.27**	**0.26**	**0.39**	**0.35**	**0.38**

Validity

Six heads of department from TU, and seven from SU, were interviewed to verify the validity of the performance scores. They were asked to give their opinion about whether the three yearly performance totals for the instructors in their department correctly reflected instructor performance. Table 4.10 shows the number and percentage of instructors whose scores were confirmed by their Heads as valid, and the number and percentage that were considered invalid.

Table 4.10 Heads of Departments' Confirmation of the Validity of the Three Yearly Performance Scores for Each of their Instructors

Instructor's Performance	TU (6 Heads)		SU (7 Heads)		Total (13 Heads)	
	Number	%	Number	%	Number	%
- Valid	18	90	18	90	36	90
- Not valid	2	10	2	10	4	10
Total	20	100	20	100	20	100

From Table 4.10, Both TU and SU Heads considered 90% of their instructors to have been scored validly.

4.3 Results of Developing the Formulae for Calculating Performance Change Scores Over a 3 Year Period composed of 4 parts as follows:

4.3.1 Developing formulae for calculating change scores

4.3.2 Test of linearity and non-linearity

4.3.3 Calculating change scores

4.3.4 Validity of change scores

4.3.1 Developing Formulae for Calculating Change Scores

It was observed from Tables 4.6 and 4.7 that mean scores for instructors rose over the three years. An objective of this project was to develop formulae for scoring individual performance change.

Instructors' performance scores may change linearly or non-linearly over the 3 year period (Academic years 1998, 1999 and 2000). Therefore, separate formulae for calculating linear and non-linear change scores were developed. The Law of Initial Values (LIV) and Floor and Ceiling Effects were considered when developing the formulae.

"According to the LIV, as scores approach a floor or a ceiling, the same external stimulus should produce smaller changes. Thus the amount of change would be a function of the starting level (i.e., the initial difference from the floor or ceiling). When approaching a ceiling, the higher scores will show a smaller increment, which will result in a negative for r_{dx} (d is change, x is initial value). When approaching a floor, the smaller scores decrease less, which also results in negative r_{dx}." (Jamieson, 1995)

The above concept implies that the instructor's performance change score depends on instructor's performance especially in the initial year. If an instructor has a high performance score in the initial year, it will be hard for him to increase his score. It shows that the increment of performance score was limited by the ceiling effect. In addition, for an instructor who had a low performance score in the initial year, his score decreased less than an instructor who had higher score. It shows that the decrease in performance score was limited by the floor effect. Thus, the performance change score is a function of change of observed scores, initial score, and floor and ceiling effects.

1) Formula for measuring change of linear performance over a 3 year period

The formula for measuring change of linear performance depends on observed score, initial score, and floor and ceiling effects as follows:

Performance change score = Change score of observed scores + Score base on initial score, floor and ceiling effects

a) Change score of Observed scores

Slope method was used to calculate the rate of change of observed scores in a 3 year period (academic years 1998, 1999 and 2000) for each instructor. It was calculated by $\frac{Y_{3i} - Y_{1i}}{3-1}$ or $\frac{Y_{3i} - Y_{1i}}{2}$ where Y_{ti} was performance score of instructor i in academic year t (t=1, 2, 3 at academic years 1998, 1999 and 2000 respectively).

In the 3 year period, calculation of slope is a simple method and the obtained scores are the same as scores calculating by least square method which is more complicated.

b) Score based on initial score, floor and ceiling effects

According to the Law of Initial Values, the magnitude of change score would be a function of the initial score. Since the research was over a 3 year period, therefore there were 2 time intervals those were: academic year 1998 to 1999 with initial score Y_{1i} and academic year 1999 to 2000 with initial score Y_{2i}. The average of both initial scores was used to represent initial score, which is $\frac{Y_{1i}+Y_{2i}}{2}$.

Initial score and floor and ceiling effects according to the concept described above were considered. The instructor who has initial score in middle point (M) especially equal 200 should have the least floor and ceiling effects, which is 0. The researcher subtracted M from the initial score that is $\frac{Y_{1i}+Y_{2i}}{2}-M$ and divided it by the difference of maximum and minimum of performance score that are 400 and 0 respectively to identify the range of $\frac{[(Y_{1i}+Y_{2i})/2]-M}{H-L}$ between -0.5 and +0.5. Then, the value of floor and ceiling effect can be multiplied with magnitude of observed change score and also combined with the observed change score to reduce the floor and ceiling effect in observed change score.

The formula was derived as shown in Equation 4.1.

$$C_i = \frac{Y_{3i}-Y_{1i}}{2} + \left[\left(\frac{\frac{Y_{1i}+Y_{2i}}{2}-M}{H-L}\right)\left|\frac{Y_{3i}-Y_{1i}}{2}\right|\right] \qquad (4.1)$$

where C_i is performance change score per academic year for instructor i

Y_{ti} is performance score of instructor i at time t (t = 1, 2, 3 at academic years 1998, 1999 and 2000)

H is maximum performance score values 400

L is minimum performance score values 0

M is middle performance score values 200

Performance change score (C_i) range from –150 (highest decreasing change) to +150 (highest increasing change), with 0 indicating no change.

Using Equation 4.1 for calculating linear change score showed that the value obtained from $\frac{[(Y_{1i}+Y_{2i})/2]-M}{H-L}$ can reduce floor or ceiling effect followed the magnitude of observed change score. It is 0 at middle point. It is more than 0 (up to +0.5) when initial score is more than middle point and less than 0 (down to –0.5)

when initial score is less than middle point. It is clear that the instructor who has the same positive change of observed scores and different initial scores, the instructor who had highest initial score is considered to have highest value of $\frac{[(Y_{1i}+Y_{2i})/2]-M}{H-L}$ that reduce the ceiling effect in positive observed change score (increase). Thus, the instructor who had highest initial score in the first year is considered to have highest positive change score (C_i). Since the instructor has a high performance score in the initial year, it will be hard for him to increase his score because of ceiling effect. (Figure 3.1)

Under floor effect, the instructor who had same negative change (decease) of observed score and different initial scores, the instructor who had lowest initial score is considered to have lowest value of $\frac{[(Y_{1i}+Y_{2i})/2]-M}{H-L}$ that reduce the floor effect in negative observed change score (increase). Thus, the instructor who had lowest initial score in the first year is considered to have highest negative change score. Because the instructor who had low performance score in the initial year, his score will be expected to decrease less than instructor who had higher score. (Figure 3.2)

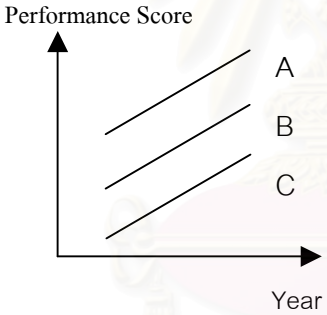


Figure 3.1 A has highest positive change (increase)

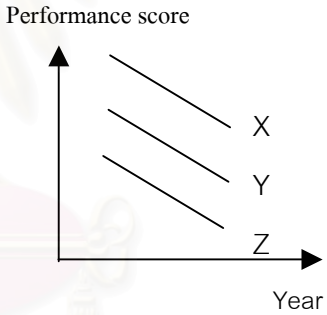


Figure 3.2 Z has highest negative change (decrease)

2) Formula for measuring change of non-linear performance over a 3 year period

From the above concept that the performance change score is a function of initial score and floor and ceiling effects, for measuring non-linear change, there are many patterns of non-linear change (i.e. parabola, quadratic, exponential etc.). From the concept of organization theory that Sudchari *et al*. (1999) developed from Schewiger *et al*. (1986), that is “Organization survives and grows when it has stability and high adaptability.” Any instructor, as a vital part of university, should improve their performance regularly. Thus, the pattern that should have the highest positive change is the pattern where performance score increases regularly; and the pattern that should have the highest negative change is the pattern where performance score decreases regularly. From both concepts, the formula for calculating non-linear change should be a function as follows:

Performance change score = f (Observed change score, initial score and floor and ceiling effects, pattern of change)

In this research, the researcher developed formula for measuring change of non-linear data over a 3 year period and the change of performance in 2 time intervals. The first time interval is academic year 1998 to academic year 1999 and the second time interval is academic year 1999 to academic year 2000.

The difference score method was used to consider the pattern of change in each period that is increase, stable, or decrease. Let Y_{ti} stand for performance score of instructor i in year t (t=1,2,3 indicating academic year 1998, 1999 and 2000). $D_{1i} = Y_{2i} - Y_{1i}$ is change score of observed scores in the first time interval (academic year 1998 to academic year 1999) and $D_{2i} = Y_{3i} - Y_{2i}$ is change score of observed scores in the second time interval (academic year 1999 to academic year 2000). The average of change score in both time intervals was used to represent observed change score, which is $\frac{D_{1i} + D_{2i}}{2}$.

According to the Law of Initial Values, the amount of change score would be a function of initial score. Since the research was over a 3 year period, therefore there were 2 time intervals. The first time interval, initial score is Y_{1i}. The second time interval, initial score is Y_{2i}. Initial score and floor and ceiling effects according to the concept described above were considered as same as linear formula. However, there were different pattern of change in non-linear (i.e. increase in a time interval and decrease in the other time interval, or decrease in a time interval and stable in the other time interval etc.). Thus, initial score and floor and ceiling effects were considered in each time interval. The consideration of initial score and floor and ceiling effects in the first time interval is $\left(\frac{Y_{1i} - M}{H - L}\right)D_{1i}$ and the second time interval is $\left(\frac{Y_{2i} - M}{H - L}\right)D_{2i}$, and both of them were considered by summing up. Then, it was multiplied with overall pattern of change (decrease or increase) that is $\frac{Y_{3i} - Y_{1i}}{|Y_{3i} - Y_{1i}|}$ for adjusting floor and ceiling effects in the case that the overall pattern of change is decrease (set $\frac{Y_{3i} - Y_{1i}}{|Y_{3i} - Y_{1i}|} = 0$, when $Y_{3i} = Y_{1i}$). The average score were calculated, which is $\frac{\left(\frac{Y_{3i} - Y_{1i}}{|Y_{3i} - Y_{1i}|}\right)\left[\left(\frac{Y_{1i} - M}{H - L}\right)D_{1i} + \left(\frac{Y_{2i} - M}{H - L}\right)D_{2i}\right]}{2}$. Then, the obtained value was combined with the observed change score to consider pattern of change and reduce the floor and ceiling effect in the observed change score. The formula was derived as shown in Equation 4.2.

$$C_i = \frac{D_{1i} + D_{2i}}{2} + \frac{\left(\frac{Y_{3i} - Y_{1i}}{|Y_{3i} - Y_{1i}|}\right)\left[\left(\frac{Y_{1i} - M}{H - L}\right)D_{1i} + \left(\frac{Y_{2i} - M}{H - L}\right)D_{2i}\right]}{2} \qquad (4.2)$$

Where $D_{1i} = Y_{2i} - Y_{1i}$ and $D_{2i} = Y_{3i} - Y_{2i}$

C_i is performance change score per academic year for instructor i

Y_{ti} is performance score of instructor i at time t (t = 1, 2, 3 at academic years 1998, 1999 and 2000)

H is maximum performance score values 400

L is minimum performance score values 0

M is middle performance score values 200

Performance change score (C_i) range from –150 (highest decreasing change) to +150 (highest increasing change), with 0 indicating no change.

Using Equation 4.2 for calculating non-linear change score, it is also found to be linear. The instructors who had same positive change of observed score, the one who had highest initial score has highest positive change score. The instructor who had same negative change of observed score, the one who had lowest initial score is considered to have highest negative change score.

For instructors who have the same initial score, the instructor who increase performance score regularly has the highest positive change score, while the instructor who decrease performance score regularly has the highest negative change score which agree with the concept of organization theory that Sudchari *et al*. (1999) developed from Schewiger *et al*. (1986).

4.3.2 Test of Linearity and Non-linearity

The performance scores of each instructor was tested for linearity by plotting graph and using analytic geometry (please see result of each instructor in Appendix D). Result of testing was summarized in Table 4.11.

Table 4.11 Percentage of TU and SU Instructors, Classify by Characteristic of Change (linear or non-linear change)

Characteristic of change	TU		SU		Total	
	Number	%	Number	%	Number	%
Linear	10	4.20	2	3.85	12	4.14
Non-linear	228	95.80	50	96.15	278	95.86
Total	238	100	52	100	290	100

From Table 4.11, The performance scores change over a 3 year period (during the academic year of 1998, 1999 and 2000) of 228 TU instructors were non-linearity (96%) as same as SU instructors, 96% were non-linearity.

4.3.3 Calculating change scores

The method of Bryk & Raudenbush (1987) was used as a validation criterion in this research. Two equations were possible, a linear model (equation 4.3) and a quadratic model (equation 4.4) according to Bryk & Raudenbush method described in section 2.5.4, chapter 2.

$$\pi_{1i} = \beta_{10} + U_{1i} \qquad \text{(4.3) (detail in p. 61-62)}$$

Where π_{1i} = the rate of change per academic year in the performance score for instructor i

β_{10} = grand mean rate of change in performance score

U_{1i} = random effect of instructor i on the rate of change in performance score

$$Q_i = \frac{\sum_{t=1}^{3} G_{it}}{3} \text{; } G_{it} = \pi_{1i} + 2\pi_{2i}(\text{AY-1998})_{it} \qquad \text{(4.4) (detail in p. 62-64)}$$

Where Q_i = the rate of change per academic year in the performance score for instructor i

π_{1i} = the instantaneous rate of change of performance score for instructor i at academic year 1998

π_{2i} = the curvature or acceleration in each growth trajectory

$(\text{AY-1998})_{it}$ = academic year (AY) of instructor i at time t minus 1998 so that $(\text{AY-1998})_{it}$ is 0, 1, 2 at academic years 1998, 1999 and 2000

Table 4.11 shows the performance scores of almost all TU and SU instructors (96%) had non-linear change. Therefore the hierarchical linear model (application of quadratic growth model) was used for calculating change scores – it makes sense to use the same formula for all change scores so that they may be compared, and in any case there were insufficient "linear" instructors for analysis. Moreover, the comparison of HLM (application of linear growth model) and HLM (application of quadratic growth model) showed the quadratic model has high reliability and high R^2 of level 1 than linear model as showed in Table 4.12.

Table 4.12 Reliability and R^2 Level 1 in Linear Model of Growth and Quadratic Model of Growth in Performance Score (unconditional model)

	TU		SU	
	HLM (linear)	HLM (quadratic)	HLM (linear)	HLM (quadratic)
Reliability of OLS Regression Coefficient Estimate				
- Initial status, π_{0i}	.965	.968	.957	.963
- Rate of change, π_{1i}	.802	.815	.741	.776
R^2 Level 1	.787	.801	.846	.867

Thus, the change scores were calculated by the researcher's formula (Equation 4.2) and hierarchical linear model (application of quadratic growth model) (Equation 4.4).

The result of calculating performance change scores of each TU and SU instructor are in **Appendix E**. The number of instructors in TU and SU who had positive change (increase), zero change (non-change), and negative change (decrease) of their performance change scores calculated by researcher's formula and Bryk & Raudenbush formula are shown in Table 4.13.

Table 4.13 Number and Percentage of TU and SU Instructor's Performance Change Scores (Negative, Zero, Positive)

University	Researcher			Bryk & Raudenbush		
	Negative	Zero	Positive	Negative	Zero	Positive
TU (n=238)	36	7	195	22	0	216
SU (n=52)	2	0	50	0	0	52
Total (n=290)	**38** *(15%)*	**7** *(2%)*	**245** *(85%)*	**22** *(8%)*	**0** *(0%)*	**268** *(92%)*

Table 4.13 showed 85% of TU and SU instructors have positive values in the change scores calculated by researcher's formula, and 92% in change score calculated by Bryk & Raudenbush formula.

Performance change scores calculated by research non-linear formula were tested the difference or analyzed variance from the groups of TU or SU instructors who had different background as shown in Table 4.14 – 4.15 for TU and Table 4.16 – 4.17 for SU.

Table 4.14 Testing the Difference of Performance Change Scores Between TU Instructors Who Have Different Sex or Internal Audit at the Faculty

Sex	**Mean**	**SD**	**t-test**
Male	11.14	12.79	0.568 (p > .05)
Female	10.24	11.57	
Internal Audit at the Faculty	**Mean**	**SD**	**t-test**
Internal Audit	13.57	12.61	3.067 (p < .05)
Not Internal Audit	8.73	11.34	

Table 4.14 shows the groups of TU instructors who have different sex did not differ significantly in performance change scores. But the group of TU instructors who worked in the faculty that did internal audit had higher mean of performance change scores than the group of instructors who worked in the faculty that did not do internal audit.

Table 4.15 Analysis of Variance of Performance Change Scores Between TU Instructors Who Have Different Backgrounds (Degree, Academic Post and Teaching Experience)

Source	**SS**	**df**	**MS**	**F**
		Degree		
Between Group	285.385	2	142.693	0.981 (p > .05)
Within Group	34189.905	235	145.489	
Total	34475.290	237		
		Academic Post		
Between Group	2009.008	3	666.669	4.827 (p < .05)
Within Group	32466.282	234	138.745	
Total	34475.290	237		

Table 4.15 (Continue)

Source	SS	df	MS	F
	Teaching Experience in Higher Education			
Between Group	4151.686	6	691.948	5.271 (p < .05)
Within Group	30323.604	231	131.271	
Total	34475.290	237		

Table 4.15 shows the three groups of TU instructors who had different graduate degree did not differ significantly in performance change scores. But the four groups of TU instructors who had different academic post differed significantly in performance change scores. The groups of instructors who were lecturers differed significantly in performance change score (mean = 13.98) from the group of instructors who were associate professor (mean = 7.64). Moreover, the seven groups of TU instructors who had different teaching experience in higher education differed significantly in performance change scores. The groups of instructors who worked less than 6 years differed significantly in performance change score (mean = 15.95) from the group of instructors who worked 26-30 years (mean = 4.74).

Table 4.16 Testing the Difference of Performance Change Scores Between SU Instructors Who Have Different Sex or Academic Post

Sex	Mean	SD	t-test
Male	22.07	14.47	0.992 (p > .05)
Female	18.25	11.22	
Academic Post	**Mean**	**SD**	**t-test**
Lecturer	21.39	13.24	1.587 (p > .05)
Assistant Professor	8.94	11.92	

Table 4.16 shows the groups of SU instructors who have different sex or academic post do not differ significantly in performance change scores.

Table 4.17 Analysis of Variance of Performance Change Scores Between SU Instructors Who Have Different Backgrounds (Degree, Teaching Experience and Faculty)

Source	SS	df	MS	F
	Degree			
Between Group	337.681	2	168.841	.940 (p > .05)
Within Group	8803.482	49	179.663	
Total	9141.163	51		
	Teaching Experience in Higher Education			
Between Group	429.787	2	214.894	1.209 (p > .05)
Within Group	8711.376	49	177.783	
Total	9141.163	51		
	Faculty			
Between Group	559.128	5	111.826	0.599 (p > .05)
Within Group	8582.035	46	186.566	
Total	9141.163	51		

Table 4.17 shows the three groups of SU instructors who had different graduate degree or teaching experience in higher education did not differ significantly in performance change scores. Moreover, the six groups of instructors who had different faculty did not differ significantly in performance change score.

4.3.4 Validity of Change Scores

1) Correlation with Bryk & Raudenbush's Change Scores

The score changes were calculated using the formula developed by the researcher. They were validated by correlating with the score changes obtained from the formula of Bryk and Raudenbush (1987). The correlation results are in Table 4.18.

Table 4.18 Pearson's Product Moment Correlation Between the Change Score from Researcher's Formula and Bryk & Raudenbush Formula (r_{XY})

Change Score	TU		SU		Total	
	n	r_{XY}	n	r_{XY}	N	r_{XY}
Non-linear	238	.984	52	.982	290	.984

Table 4.18 shows the correlations ranged between .982 - .984.

2) Seeking Opinions of Heads of Department and Instructors

Further validation was obtained by seeking the opinions of heads of department and instructors, who were asked to compare the change scores obtained from the researcher's formula with those obtained from Bryk & Raudenbush's. (See Table 4.19)

Table 4.19 Opinions of Heads of Department and Instructors about the Validity of Change Score Calculated by the Researcher Formula (F1) and Bryk&Raudenbush's Formula (F2). F1 > F2 means F1 was considered better; F1 = F2 means neither score was considered better; F1 < F2 means F2 was considered better;

Department, University	Evaluatee	Evaluator					
		Head of department			Instructor (Themselves)		
		F1 > F2	F1 = F2	F1 < F2	F1 > F2	F1 = F2	F1 < F2
Department A, TU	TU1	✓*			✓*		
	TU2	✓*			✓*		
	TU3	✓*			✓*		
	TU4	✓				✓	
Department B, TU	TU5		✓*			✓*	
	TU6	✓*			✓*		
	TU7			✓*			✓*
Department C, TU	TU8			✓		✓	
	TU9			✓	✓		
	TU10			✓*			✓*
Department D TU	TU11	✓*			✓*		
	TU12	✓*			✓*		
	TU13	✓				✓	
	TU14	✓*			✓*		
Department E, TU	TU15	✓*			✓*		
	TU16	✓*			✓*		
Department F, TU	TU17		✓		✓		
	TU18		✓*			✓*	
	TU19		✓		✓		
	TU20		✓*			✓*	
Department A, SU	SU1	✓*			✓*		
	SU2	✓*			✓*		
	SU3	✓*			✓*		
	SU4	✓				✓	
Department B, SU	SU5			✓		✓	
	SU6	✓*			✓*		

Table 4.19 (Continue)

Department, University	Evaluatee	Evaluator					
		Head of department			Instructor (Themselves)		
		F1 > F2	F1 = F2	F1 < F2	F1 > F2	F1 = F2	F1 < F2
Department C, SU	SU7	✓*			✓*		
	SU8	✓*			✓*		
	SU9	✓*			✓*		
	SU10	✓*			✓*		
Department D, SU	SU11		✓		✓		
	SU12		✓		✓		
Department E, SU	SU13	✓*			✓*		
	SU14	✓*			✓*		
Department F, SU	SU15	✓*			✓*		
	SU16			✓		✓	
	SU17			✓*			✓*
	SU18			✓	✓		
Department G, SU	SU19	✓				✓	
	SU20	✓					✓
Sum		**25**	**7**	**8**	**26**	**10**	**4**

Note *Head of department and Instructor have the same opinion

A summary of the opinions of heads of department and instructors from Table 4.19 is presented in Table 4.20.

Table 4.20 Opinions of Heads of Department and Instructors about the Validity of Change Score Calculated by the Researcher Formula (F1) and Bryk&Raudenbush's Formula (F2)

Head of Department	Instructor (n = 40)			
(n = 13)	F1 > F2	F1 = F2	F1 < F2	**Total**
F1 > F2	***20 (50%)***	4 (10%)	1 (2.5%)	**25 (62.5%)**
F1 = F2	4 (10%)	***3 (7.5%)***	0 (0%)	**7 (17.5%)**
F1 < F2	2 (5%)	3 (7.5%)	***3 (7.5%)***	**8 (20%)**
Total	**26 (65%)**	**10 (25%)**	**4 (10%)**	**40 (100%)**

From Table 4.20 it can be seen that there was a large degree of agreement between heads of department and instructors. 65% of the instructors considered scores from the researcher's formula to be better than those from Bryk & Raudenbush's formula, compared with 10% who thought Bryk and Raudenbush's better. Heads of department considered 62.5% of the researcher's scores to be better, and only 20% of Bryk and Raudenbush's to be better.

Chapter 5
Discussion and Conclusion

The study had two main objectives: 1) to develop a methodology to measure university instructors' performance and 2) to develop the formulae for calculating performance change scores over a 3 year period.

Following a literature review and interviews with senior academic staff from Thammasat University and Saint John's University, 35 subtasks, grouped under six principal tasks, were proposed for the scoring of university instructors' performance. The subtasks were endorsed as being appropriate by further samples of senior staff from both universities. Additional samples of senior and experienced staff were consulted to determine the weight each task and subtask should contribute to instructor performance scores at each university. Data for fourteen of the subtasks could be obtained from existing university documents; a self-report questionnaire was developed for instructors to give data for the remaining 21 subtasks.

Data was collected for three academic years 1998, 1999 and 2000 for a total of 290 instructors from both universities. A performance score was obtained by multiplying the 35 subtask scores by the developed weighting and summing the weighted scores. Performance scores were reliable and the relationship between subtasks and total score was stable across the years within each university. Heads of department considered the scores to be valid measures of 90% of the instructors in their departments.

The formulae for calculating change scores were developed, one for linear change and the other for non-linear change. Since almost all the instructors (96%) for whom data were collected exhibited non-linear change, only the non-linear formula could be evaluated. Performance change scores calculated using the researcher's formula were validated by calculating the correlation between them and change scores obtained from the Hierarchical Linear Model (application of quadratic growth model) of Bryk & Raudenbush (1987). The correlation was 0.98 for both universities. Further validation was obtained by asking heads of department and instructors to compare the change scores obtained from the researcher's formula with those obtained from Bryk & Raudenbush's method. 65% of the instructors considered their score from the researcher's formula to be better than that from Bryk & Raudenbush's formula, compared with 10% who thought Bryk and Raudenbush's better. Heads of department

considered 62.5% of the researcher's scores to be better, and only 20% of Bryk and Raudenbush's to be better. There was therefore evidence that change scores calculated using the researcher's non-linear formula were valid.

Discussion

Discussion of results in relation to objectives

1. Discussion of results in relation to **objective 1** "to develop a methodology to measure university instructors' performance scores"

This objective composes 3 sub-objectives that are to identify the major tasks of the instructor, to identify weighting of tasks and subtasks and to calculate performance scores.

Table 4.3 shows senior and experienced staff at both SU and TU confirmed that instructors' duties could be divided into six principal tasks consisting of 35 subtasks that were appropriate at both universities. The six principal tasks confirmed were the same as tasks from the literature review. Almost all subtasks identified were related to subtasks obtained from the literature review and considered both quantity and quality.

Table 4.4 shows considerable agreement between the two universities in the weight with which each task and subtask should contribute to total performance scores. Both universities considered teaching to have the highest weight, and remarkably they both gave it an identical number of points. The differences that did exist are unsurprising given the two universities' different missions and backgrounds. For example, TU gave eight points more weight to research and academic publications than SU, whilst SU gave six more points than TU to student advice. However, there was a difference of no more than one point between the two universities for 25 of the 35 subtasks, and only one of the subtasks, the number of texts or academic books published, had a difference of more than two points (with TU unsurprisingly giving it 2.2 more points than SU). Nevertheless, the differences were sufficient for it to make sense to use different weightings for calculating performance scores at TU and SU.

The lowest scoring subtask (as a proportion of the maximum score available) at both universities in the final year was 3.2 (number of research projects contributed to), closely followed by 3.3 (number of text or other academic books written). It may be that the criteria for scoring these subtasks were too severe, though in both cases there was at least one instructor at both universities who received the maximum score possible, so

it was clearly possible to do well and the low scores may instead reflect generally poor performance in these subtasks at both universities. It is not surprising that the highest scoring subtask at SU was 1.11 since this private university considers student satisfaction extremely important, and instructors are likely to be well motivated to achieve it.

The Cronbach's alpha figures in Table 4.8 indicate the degree to which the subtasks measured the same construct. All the figures are around 0.8, indicating that the scores had acceptable internal reliability each year at each university, even though SU's were marginally more reliable.

When subtasks were ranked each year at each university by their correlation with the sum of all other subtasks (Table 4.9), Spearman's Rho correlation between the rankings ranged from 0.88 to 0.96 at TU, and 0.88 to 0.90 at SU. These high correlations indicate considerable stability across the years within each university. Similar Spearman's Rho correlations were obtained when the subtasks whose data came from documents, and the subtasks whose data came from questionnaires, were ranked separately, so this inter-year stability does not seem to be a result of the fact that questionnaire data was collected for all three years simultaneously.

When the subtasks were ranked by the three-yearly average of their correlations at each university, Spearman's Rho between TU and SU was 0.42. Although this correlation is statistically significant at the 0.05 level, it is substantially less than the within-university correlations, implying that there were differences between the universities in the relationships between the subtasks and the total. Incidentally, this difference is not due to the different weightings given at the different universities: the Spearman's Rho correlation is also 0.42 for the unweighted data.

Subtask 3.2 was one of the few subtasks which had a very low correlation with the total of the other subtasks at both universities. However, this was probably because most instructors scored lowly on this subtask (Tables 4.6 and 4.7), which gave little scope for scores to vary and therefore correlate with the total.

If the three correlations for each subtask at each university in Table 4.9) are averaged, five subtasks had average correlations of more than 0.4 at both universities. These were 1.4 (use of appropriate teaching methods), 1.5 (providing stimulating activities and environment), 1.12 (the use the instructor makes of evaluations of his performance), 3.5 (quality of academic work) and 4.4 (the degree to which the instructor's service to the community fit into the faculty plan or policy). These five

subtasks were therefore quite strongly related to instructor performance at both universities.

There were three subtasks for which the average correlation was 0.4 (or more) higher at SU than at TU. These were 1.3 (teaching load), 1.9 (examination administered on time) and 1.10 (grade reported to faculty on time). Subtasks 1.9 and 1.10 in fact had the highest correlations of any subtask at SU, and amongst the lowest for TU. These differences are not surprising, however, given the differences in mission and organization between the two universities. There were few subtasks for which the average correlation was higher at TU than at SU. The largest differences (0.18 and 0.15 respectively) in TU's favour were for subtasks 3.2 (number of research projects contributed to) and 3.4 (number of teaching materials written). Subtask 3.2 was flagged above for having very weak correlations at both universities (an average of -0.09 at SU and +0.09 at TU), so it is worth noting that the subtask for which the difference in favour of TU was largest was hardly related to overall performance scores at either university. Nevertheless, the two subtasks for which the difference (however marginal) in favour of TU was largest were both part of the Research and Academic Publications task, which was to be expected given TU's mission.

Both Cronbach's alpha and the difference between subtask-total correlations at the two universities suggest that the performance scores were slightly more reliable for SU than for TU. This might reflect the fact that instructors at TU have more autonomy, fostering greater variation in the subtasks that individual instructors apply themselves to and, consequently, less unidimensionality. The fact that the inter-year correlations for TU were high (marginally higher than for SU, in fact) suggests that the slightly lower reliability is due to scores being consistently slightly less unidimensional at TU, rather than due to random error. However, the difference in reliability between the two universities is not large, and the scores have adequate reliability at both universities.

Table 4.10, shows that both TU and SU Heads considered 90% of their instructors to have been scored validly, providing evidence that the performance scores produced had a high degree of validity at both universities. Noticeably, the scores of their instructors, which decrease in the following year, were considered to be valid.

Thus the performance scores were reliable and valid at both universities, though some subtasks had different weights and the relationship between the subtasks and total score was different.

2. Discussion of results in relation to **objective 2** "to develop the formulae for calculating performance change scores over a 3 year period"

This objective composes 2 sub-objectives that are to develop the formulae for calculating performance change scores and to validate change scores using the formulae developed by researcher.

Table 4.11 shows almost all TU and SU instructors have non-linear change, therefore, only non-linear formula were validate.

Table 4.13 shows almost all TU and SU instructors had positive change scores that mean almost all instructors improved their performance when their university introduced quality assurance (during academic years 1998-2000). The result obtained related to Newton (1999) 's research that 73% of academic managers agree that the quality system had led to improvement in quality for staff, and also improve the quality of teaching (Dill, 2000, Fourie & Alt, 2000).

Table 4.14 shows Pearson's product moment correlation between the change scores from linear formula of researcher and those of Bryk & Raudenbush (1987) ranged between .982-.984, which were very high positive correlation. In Bryk & Raudenbush method, Empirical Bayes (EB) was used to estimate change score for each instructors. EB provides a composite estimator, π_i^*, which rate of change of subject i estimated by mean of ordinary least square (OLS), $\hat{\pi}_i$, based on the repeated measurement for that subject and reliability of the ordinary least square estimate, W_i, as show in Equation $\pi_i^* = W_i\hat{\pi}_i + (1-W_i)\hat{\pi}_i$ (Bryk & Raudenbush, 1987). If reliability equals 1, the EB estimator will have the same value as OLS estimator. In researcher's formulae, a change score is observed change score that consider pattern of change and reduce the floor and ceiling effects. Observed change scores were calculated by $\frac{D_{1i} + D_{2i}}{2} = \frac{Y_{3i} - Y_{1i}}{2}$. For 3 year periods, they are the same scores as calculated by OLS. Table 4.12 shows reliability of rate of change valued .815 for TU and .776 for SU (quite high reliability), therefore, scores obtained from both methods had high correlation because they had the same common factor for calculating change score.

Table 4.16 shows 65% of instructors and 62.5% of heads of department considered scores from the researcher's formula to be better than those from Bryk & Raudenbush's formula. Noticeably, it does not differ between the considerations of change scores of instructors who have different pattern of change. Moreover, the

obtained formula is easy to calculate and translate the meaning without impact of sample size.

Thus, the change score obtained from researcher's non-liner formula were valid and easy to use.

Discussion of results in relation to hypotheses

1. Discussion of results in relation to **hypothesis 1** "A methodology for scoring university instructors' performance that covers all their major tasks and considers both quantity and quality by multiplying the scores of subtasks developed with the weighting obtained from their university's experts and summing the weighted scores, should be an appropriate methodology."

Most of the studies in instructors' performance emphasize only quantitative, objective criteria, for example the number of hours worked, e.g. Umpuang (1985), Pittayanuwat et al. (1981), Boonying (1986) and Punsuwan (1994). Moreover, some research studied only teaching, such as the research by Buakam (1997) and Boondeekul (1998), even though instructors' tasks are much broader than teaching alone. But in this research, Table 4.3 shows the six principal tasks obtained from the literature review were confirmed. Moreover, 35 subtasks selected by senior and experienced staff were confirmed. Almost all subtasks related to literature reviews and covered both quantity and quality. Moreover, Table 4.9 shows the correlation of all the subtasks each year was around 0.3 for TU and 0.4 for SU. Almost all correlations were positive, and none were strongly negative. The variable has a low correlation with other variables and a moderate correlation with the concept should be selected (Johnstone, 1981). Thus, subtasks developed were available.

Table 4.4 shows both universities considered teaching as the most important task (the same weight). The other different weighting of tasks showed the different mission and background of TU and SU. TU gave eight points more weight to research and academic publications than SU, whilst SU gave six more points than TU to student advice. Because TU is a leading public university that composed of many doctoral degree instructors and offers graduate programs up to doctoral degrees (TU, 2002), in the other hand, SU is private university emphasized in ISO9002 that student service is the most important (SU, 2002). Expert judgment can determine weights (Johnstone, 1981). Thus, the weightings obtained were believable.

Moreover, Table 4.8 shows reliability of performance scores around 0.8, indicating that the scores have acceptable reliability and Table 4.10 shows the scores had high validity. Thus, the methodology developed was an appropriate methodology.

2. Discussion of results in relation to **hypothesis 2** " The performance scores change over a 3 year period, should be non-linear."

Table 4.11 shows almost all TU and SU instructors have non-linear change. The result agrees with many researchers e.g. Tangsakulruanglai (1998), Khamlan (1997), Chan et al. (2000), Silverstein and Long (1998), Wijitwanna (2000) who showed the variables that they studied had non-linear or quadratic change.

Moreover, the floor and ceiling effects may influence performance score especially instructors who had higher scores in the initial year. The higher scores will show a smaller increment (Jamieson, 1995). Thus, it is difficult for them to increase their performance consistently which makes them have non-linear change.

Instructors have academic autonomy, especially in public universities, thus the increase or decrease in their performance depended on the individual. Furthermore, environment change always continually happens in their organization. It may influence the change in the performance of instructors (Prapavanon et al, 1998). Therefore, the performance scores of instructors do not change consistently, but in a non-linear manner.

Suggestions

Weakness and problems in this research

Even though the researcher received very good cooperation from TU and SU staff, the data collection still had problems in some faculties because they did not have systematic documents especially on instructor's work and some faculties refused to give data. Moreover, the response rate in some faculties was quite low when compared with other faculties and also the response rate in TU is lower than SU. This illustrates usefulness for collecting data in university was the support from senior staff especially President, Vice-President and Dean.

Almost all instructors in TU and SU showed non-linear change, thus only non-linear formula were validated. In other universities, if almost all instructors in their faculty or university have linear change, it may be necessary to use linear formula for calculating performance change score because it was developed from the same concept with non-linear formula. But those scores should be validated before using the results.

Suggestions for practical implication

The methodology for measuring instructors' performance scores developed is suitable for measuring instructors in higher education institutions but the weighting of tasks and subtasks should be identified by their expert for each separate institution. The formula developed is suitable for calculating performance change score for any case which collects data in 3 year periods or 3 time points.

The performance scores and performance change scores obtained are useful for administration such as consideration for promotion, instructor evaluation and instructor development. Since it is easy to use them to compare the performance of their instructors and the change in them.

Moreover, the methodology for measuring instructors' performance developed was accepted for publishing in international journal so it is a good methodology that should be consider for measuring instructor's performance scores.

Suggestions for further work

In this research, the results showed that many instructors had improvement in their performance after the introduction of quality assurance in university. But in this research, it is difficult to summarize exactly that they have improved because of quality assurance.

For further work, therefore, it should have universities that didn't introduce quality assurance to be a control group for comparison in performance change scores between universities that introduced and didn't introduce quality assurance. Moreover, it should have the studies about what factors or variables that influence an individual to improve their performance and the appropriate weighting of them. The results will be useful for administration and the development of university because instructors are a vital part of a university and their performance reflect quality of their university.

Conclusion

The first objective of the research was to develop a methodology for scoring university instructors' performance, which covered all their major tasks and considered both quantity and quality. That objective was met. It was found that performance could be divided into six principal tasks and 35 subtasks that were appropriate for two very different Thai universities. Different weighting factors for tasks and subtasks needed to be applied for the different universities, however, and the relationship between subtasks and overall scores differed between the universities. This was expected given the universities' different missions and backgrounds, and the scores were reliable and considered valid at both universities.

The second objective of the research was to develop the formulae for calculating performance change scores over a 3 year period. It was found that almost all instructors in both university has non-linear change, therefore, only non-linear formula were validated in this research. However, the formula has high validity for the very different universities, moreover, it is easy to calculate and translate the meaning without impact of sample size. Thus, the non-linear formula developed was suitable for calculating performance change score for each instructor over a 3 year period.

References

Thai (These references are written in Thai)

กุลนภา ฟู่เจริญ. 2531. หน้าที่ของอาจารย์ในมหาวิทยาลัย. บทบาทอาจารย์มหาวิทยาลัย, หน้า 5-13. ขอนแก่น: มหาวิทยาลัยขอนแก่น.

ช่อลัดดา ขวัญเมือง. 2524. การศึกษาบทบาทของอาจารย์ที่ปรึกษาในวิทยาลัยครูภาคเหนือ. วิทยานิพนธ์ศิลปศาสตรมหาบัณฑิต ภาควิชาจิตวิทยาการศึกษาและแนะแนว มหาวิทยาลัยธรรมศาสตร์.

ชวลิต ประภวานนท์และคณะ. 2541. พฤติกรรมองค์การ. กรุงเทพฯ: ธีระฟิล์มและไซเท็กซ์.

ทบวงมหาวิทยาลัย. 2535. ประมวลประกาศ ระเบียบ หลักเกณฑ์ในส่วนที่เกี่ยวข้องกับสถาบันอุดมศึกษาเอกชน. กรุงเทพฯ : กองสถาบันอุดมศึกษาเอกชน.

ทองใบ สุดชารี และคณะ.2542. ทฤษฎีองค์การ : วิเคราะห์แนวความคิด ทฤษฎีและการประยุกต์. อุบลราชธานี: คณะวิทยาการจัดการ สถาบันราชภัฏอุบลราชธานี.

นุชรินทร์ บุญยิ่ง. 2529. การพัฒนารูปแบบการวิเคราะห์ปริมาณงานอาจารย์วิทยาลัยครู. วิทยานิพนธ์ครุศาสตรมหาบัณฑิต ภาควิชาวิจัยการศึกษา จุฬาลงกรณ์มหาวิทยาลัย.

นงเยาว์ บัวคำ. 2540. ภาระงานสอนของอาจารย์คณะศึกษาศาสตร์ ปีการศึกษา 2540. เชียงใหม่: คณะศึกษาศาสตร์ มหาวิทยาลัยเชียงใหม่.

ประดิษฐ์ มีสุข. 2543. การประเมินประสิทธิภาพการสอนของอาจารย์. คู่มือการสอน, หน้า 12-20. สงขลา: มหาวิทยาลัยทักษิณ.

พงษ์ศักดิ์ วรรณไกรโรจน์และคณะ. 2538. การเพิ่มศักยภาพและการประเมินบุคลากร. ในรายงานการสัมมนาการเพิ่มศักยภาพและการประเมินคุณภาพบุคลากรมหาวิทยาลัย, หน้า 47-53. 8-10 กันยายน ณ โฮมพุเตย รีสอร์ท จังหวัดกาญจนบุรี.

พวงแก้ว ปุณยกนก และนพรัตน์ ทิสยากร. 2526. การประเมินประสิทธิภาพการสอนของอาจารย์จุฬาลงกรณ์มหาวิทยาลัย. รายงานการวิจัย ฝ่ายวิชาการ จุฬาลงกรณ์มหาวิทยาลัย.

พรชุลี อาชวอำรุง. 2537. บทบาทอาจารย์ที่ปรึกษาวิชาการเพื่อการพัฒนานิสิตนักศึกษาในมหาวิทยาลัยของรัฐ. วารสารครุศาสตร์. (มกราคม-มีนาคม): 89-96.

พรเลิศ อ่วมพ่วง. 2528. การวิเคราะห์ภาระงานของอาจารย์มหาวิทยาลัยศิลปากร วิทยาเขตพระราชวังสนามจันทร์. วิทยานิพนธ์ปริญญาครุศาสตรมหาบัณฑิต ภาควิชาอุดมศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

ไพฑูรย์ ปุ่นสุวรรณ (2537). ภาระงานของอาจารย์วิทยาลัยครู : การศึกษาเฉพาะกรณีวิทยาลัยครูสุราษฎร์ธานี. วิทยานิพนธ์ครุศาสตรดุษฎีบัณฑิต ภาควิชาอุดมศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

ไพฑูรย์ สินลารัตน์. 2530. การอุดมศึกษากับสังคมไทย. กรุงเทพฯ : โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.

ภัทรฤดี สุพัฒโสภณ. 2535. การวิเคราะห์บทบาทของอาจารย์มหาวิทยาลัยสุโขทัยธรรมาธิราช. วิทยานิพนธ์ครุศาสตรมหาบัณฑิต ภาควิชาอุดมศึกษา จุฬาลงกรณ์มหาวิทยาลัย.

มหาวิทยาลัยเซนต์จอห์น. 2543. รายงานประเมินตนเองของมหาวิทยาลัยเซนต์จอห์น. (อัดสำเนาเย็บเล่ม)

มหาวิทยาลัยเซนต์จอห์น. 2541. มหาวิทยาลัยเซนต์จอห์น: มหาวิทยาลัยเอกชนแห่งแรกที่ได้รับมาตรฐาน ISO9002. กรุงเทพฯ: มหาวิทยาลัยเซนต์จอห์น.

มหาวิทยาลัยเซนต์จอห์น. 2540. คู่มืออาจารย์ประจำ มหาวิทยาลัยเซนต์จอห์น. กรุงเทพฯ: แผนกบุคคล.

มหาวิทยาลัยทักษิณ. 2542. คู่มือโครงสร้างระบบประกันคุณภาพการศึกษามหาวิทยาลัยทักษิณ. สงขลา: มหาวิทยาลัยทักษิณ.

มหาวิทยาลัยธรรมศาสตร์. 2543. รายงานประเมินตนเองมหาวิทยาลัยธรรมศาสตร์. (อัดสำเนาเย็บเล่ม)

มหาวิทยาลัยอุบลราชธานี. 2544. เกณฑ์คำนวณภาระงานเชิงปริมาณของอาจารย์คณะศิลปศาสตร์ มหาวิทยาลัยอุบลราชธานี. (เอกสารอัดสำเนา)

มาณี ไชยธีรานุวัฒศิริ. 2533. การประเมินผู้บริหารและอาจารย์ : สิ่งส่งเสริมความสำเร็จของสถาบันอุดมศึกษา. อุดมศึกษา : ทางสู่ความสำเร็จ, หน้า 13-22. กรุงเทพฯ: ภาควิชาอุดมศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

วรรณี จำนงรักษ์. 2532. บทบาทของอาจารย์ที่ปรึกษาตามความคิดเห็นของอาจารย์ที่ปรึกษาและนักศึกษาในวิทยาลัยสังกัดกรมอาชีวศึกษาจังหวัดสุราษฎร์ธานี. วิทยานิพนธ์ศิลปศาสตรมหาบัณฑิต ภาควิชาจิตวิทยาการศึกษาและแนะแนว มหาวิทยาลัยธรรมศาสตร์.

วันชัย ศิริชนะ. 2537. การพัฒนารูปแบบการประกันคุณภาพการศึกษาระดับอุดมศึกษาสำหรับสถาบันอุดมศึกษาในสังกัดทบวงมหาวิทยาลัย. วิทยานิพนธ์ปริญญาดุษฎีบัณฑิต สาขาวิชาอุดมศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

วัลลภา เทพหัสดิน ณ อยุธยา. 2543a. ความเป็นครูและจรรยาบรรณวิชาชีพอาจารย์ระดับอุดมศึกษา. กรุงเทพฯ: คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

วิไลวรรณ พิธิยานุวัฒน์ และคณะ. 2524. การวิเคราะห์ปริมาณงานอาจารย์ประจำจุฬาลงกรณ์มหาวิทยาลัย. กรุงเทพฯ: จุฬาลงกรณ์มหาวิทยาลัย.

วีระศักดิ์ คำล้าน. 2540. การวิเคราะห์การเปลี่ยนแปลงในระยะยาว ของผลสัมฤทธิ์ทางการเรียนคำศัพท์ภาษาอังกฤษ : การประยุกต์ใช้โมเดลเชิงเส้นพหุระดับ. วิทยานิพนธ์ครุศาสตรมหาบัณฑิต ภาควิชาวิจัยการศึกษา จุฬาลงกรณ์มหาวิทยาลัย.

ศาณี สุวรรณพัฒน์. 2533. คุณภาพอาจารย์ในสถาบันอุดมศึกษา. อุดมศึกษา : ทางสู่ความสำเร็จ, หน้า 67-81. กรุงเทพฯ: ภาควิชาอุดมศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.

ศักดิ์ชาย เพชรช่วย. 2541. การพัฒนาตัวบ่งชี้รวมคุณภาพการศึกษาของคณะครุศาสตร์ในสถาบันราชภัฏ. วิทยานิพนธ์ครุศาสตรมหาบัณฑิต ภาควิชาวิจัยการศึกษา จุฬาลงกรณ์มหาวิทยาลัย.

สมจิตต์ สุพรรณทัศน์และคณะ. 2527. การประเมินผลการปฏิบัติงานด้านการบริการของอาจารย์มหาวิทยาลัย. ในรายงานการสัมมนาทางวิชาการเรื่องการประเมินผลการปฏิบัติงานของอาจารย์มหาวิทยาลัย, หน้า 37-46. 16-17 พฤษภาคม ณ คณะแพทยศาสตร์ศิริราชพยาบาล.

สมถวิล วิจิตรวรรณา. 2543. การเปรียบเทียบประสิทธิภาพของโมเดลโค้งพัฒนาการที่มีตัวแปรแฝงโมเดลพหุระดับ และโมเดลกึ่งซิมเพลกซ์ ในการวัดการเปลี่ยนแปลงระยะยาวชนิดตัวแปรเดี่ยวและตัวแปรพหุ. วิทยานิพนธ์ครุศาสตรดุษฎีบัณฑิต ภาควิชาวิจัยการศึกษา จุฬาลงกรณ์มหาวิทยาลัย.

สมบุญ บุญดีกุล. 2541. การวิเคราะห์ภาระงานสอนของอาจารย์จุฬาลงกรณ์มหาวิทยาลัย. กรุงเทพฯ: งานวิจัยสถาบัน กองแผนงาน ฝ่ายวางแผนและพัฒนา จุฬาลงกรณ์มหาวิทยาลัย.

สัมพันธ์ พันธุ์พฤกษ์. 2538. ความตรงและความเที่ยงของการประเมินอาจารย์มหาวิทยาลัยโดยเพื่อนร่วมงาน : การศึกษาเฉพาะกรณี. วิทยานิพนธ์ครุศาสตรดุษฎีบัณฑิต ภาควิชาวิจัยการศึกษา จุฬาลงกรณ์มหาวิทยาลัย.

สำเนาว์ ขจรศิลป์. 2530. การศึกษาระบบอาจารย์ที่ปรึกษาด้านวิชาการระดับปริญญาตรี ในมหาวิทยาลัยเกษตรศาสตร์ มหาวิทยาลัยศิลปากร และมหาวิทยาลัยสงขลานครินทร์. รายงานการวิจัย มหาวิทยาลัยเกษตรศาสตร์.

สำนักมาตรฐานอุดมศึกษา. 2539. คู่มือการพัฒนาระบบการประกันคุณภาพการศึกษาสำหรับสถาบันอุดมศึกษา. กรุงเทพฯ: ทบวงมหาวิทยาลัย.

สุวรรณา ทรัพย์บำรุง. 2537. การวิเคราะห์บทบาทของอาจารย์พยาบาลในวิทยาลัยพยาบาลทหารตำรวจ. วิทยานิพนธ์ปริญญาครุศาสตรมหาบัณฑิต ภาควิชาอุดมศึกษา บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย.

สุภารัตน์ เรือจันทึก. 2542. การวิเคราะห์การเปลี่ยนแปลงระยะยาวของผลสัมฤทธิ์ทางการเรียนสิ่งแวดล้อม : การเปรียบเทียบผลการวิเคราะห์ด้วยโมเดลเอชแอลเอ็ม และโมเดลลิสเรล. วิทยานิพนธ์ครุศาสตรมหาบัณฑิต ภาควิชาวิจัยการศึกษา จุฬาลงกรณ์มหาวิทยาลัย.

อดุลย์ วิริยเวชกุล. 2541. การประกันคุณภาพทางวิชาการในมหาวิทยาลัยไทย. กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยมหิดล.

อาทิตยา ดวงมณี. 2540. การพัฒนาตัวบ่งชี้รวมสำหรับความเป็นเลิศทางวิชาการของสาขาวิชาทางวิจัยการศึกษาในมหาวิทยาลัยของรัฐ. วิทยานิพนธ์ครุศาสตรมหาบัณฑิต ภาควิชาวิจัยการศึกษา จุฬาลงกรณ์มหาวิทยาลัย.

อุทัย บุญประเสริฐ และคณะ. 2542. อาจารย์มหาวิทยาลัย. รายงานวิจัยเรื่องการเตรียมความพร้อมของมหาวิทยาลัยของรัฐเพื่อเป็นมหาวิทยาลัยในกำกับ. สำนักงานกองทุนสนับสนุนการวิจัย.

อุทุมพร จามรมาน. 2541. การตรวจสอบผลการประกันคุณภาพการอุดมศึกษา. คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย. (เอกสารอัดสำเนา)

อุทุมพร จามรมาน. 2543b. การประเมินการสอน. คู่มือการสอน, หน้า 21-29. สงขลา: มหาวิทยาลัยทักษิณ.

อุทุมพร จามรมาน. 2543b. การประกันคุณภาพระดับอุดมศึกษา. กรุงเทพฯ: โรงพิมพ์ฟันนี่.

อุทุมพร จามรมานและคณะ. 2544. ดัชนี เกณฑ์และวิธีการประเมินคุณภาพการศึกษาภายในและภายนอกสำหรับอุดมศึกษาไทย. รายงานการวิจัย ทุนอุดหนุนจากสำนักงานกองทุนสนับสนุนการวิจัย.

อุทุมพร จามรมาน. 2544. วิธีทำการประกันคุณภาพการศึกษาของสถาบันอุดมศึกษา. กรุงเทพฯ: โรงพิมพ์ฟันนี่.

English

Baer, D and et al. 1996. Faculty Workload Study. The University of Western Ontario [online]. Available from: http://www.uwo.ca/western/workload/workload.html [2000, May 18].

Bijleveld, C.J.H. and et al. 1998. Longitudinal data analysis : Designs, models and methods. London: Sage Publications.

Braun, H.I. and et al. 1983. Empirical Bayes estimation of coefficients in the general linear model from data of deficient rank. Psychometrika. 48(2): 171-181.

Braun, H.I. 1988. Empirical Bayes methods: A tool for exploratory analysis. Edit by Bock, R.D. Multilevel analysis of educational data. San Diego : Academic Press, Inc.: 19-55.

Brennan, J., Frederiks, M. and Shah, T. 1997. Improving the quality of education the impact of quality assessment on institutions. Hampshire: Hobbs the Printers Ltd.

Bryk, A.S. and Raudenbush, S.W. 1987. Application of hierarchical linear models to assessing change. Psychological bulletin, 10(1): 147-158.

Bryk, A.S. and Raudenbush, S.W. 1992. Hierarchical linear models: Applications and data analysis methods. London: Sage Publications.

Carroll, M.L. 1997. Self-audits of quality in academic and service department: A practical experience. Assessment & evaluation in higher education. 22(2): 225-231.

Cave, M. and et al. 1997. The use of performance indicators in higher education. London: Jessica Kingsley.

Chan, and et al. 2000. Modeling intraindividual change in children's social skills at home and at school : A multivariate latent growth approach to understanding between-settings differences in children's social skill development. Multivariate behavioral research. 35(3): 365-396.

CHEA, 1998. CHEA Recognition. [online] Available from: www.chea.org [2000, July 15].

CHEA, 2000a. Council for Higher Education Accreditation. [online] Available from: www.chea.org [2000, July 15].

CHEA, 2000b. Council for Higher Education Accreditation: Directories. [online] Available from: www.chea.org [2000, July 15].

Chou, C., Bentler, P.M. and Pentz, M. 1998. Comparisons of two statistical approaches to study growth curves : The multilevel model and the latent curve analysis. Structural equation modeling. 5(3): 247-266.

Collins, L.M. and Cliff, N. 1990. Using the longitudinal Guttman simplex as a basis for measuring growth. Psychological bulletin. 108(1): 128-134.

Cryer, P. 1993. Preparing for quality assessment and audit. Sheffield: CVCP/USDU.

Dill, D.D., Massy, W.F. and Williams, P.R. 1996. Accreditation & academic quality assurance. Change. 28: 16-24.

Dill, D.D. 2000. Designing academic audit: lessons learned in Europe and Asia. Quality in higher Education. 6(3): 187-207.

Donaldson, J.1994. Quality assurance in Scotland. In A. Craft (ed.), International developments in assuring quality in higher education, p.102-111. London: The Falmer Press.

Duhs, L.A. 2000. Performance indicators and research performance of academic groups. [online] Available from: http://www.deetya.gov.au/highered/hereview/submissions/submissions/d/duhs3.htm [2000, July 15].

Embretson, S.E. 1991. A multidimensional latent trait model for measuring learning and change. Psychometrika. 56(3): 495-515.

FEFCW. 1999. Quality assurance handbook. [online] Available from: http://www.wfc.ac.uk/fefew/Index.html [2000, July 15].

Ferguson, G.A. 1971. Statistical analysis in psychology and education. Tokyo: Mcgraw-Hill Kogakusha, Ltd.

Fourie & Alt. 2000. Challenges to sustaining and enhancing quality of teaching and learning in South African universities. Quality in higher education. 6(2): 115-124.

Frazer, M. 1991. Quality assurance in higher education. In A. Craft, Quality assurance in higher education, p.11-27. London: The Falmer Press.

French, N. 1997. Financing higher education in Hong Kong: Outline of the UGC funding methodology. Hong Kong: UGC.

Fuller, G. 1967. Analytic geometry. Massachusetts: Addison-Wesley Publishing company.

Gupta, C.B. 1993. An introduction to statistical methods. New Delhi : Vikas.

Gottman, J.M. 1995. The analysis of change. New Jersey: Lawrence Erlbaum Associates, Publishers.

Howell, D.C. 1992. Statistical methods for psychology. California: Duxbury Press.

Hughes, T., Williams, T., Ryall, P. 2000. It is not what you achieve it is the way you achieve it. Total quality management. 11(3): 329-340.

Ibid, K.H. and et al. 2000. Socio-structural change and educational development in the Asia Pacific Region. Houndmills: Macmillan.

Jamieson, J. 1995. Measurement of change and the law of initial values: a computer simulation study. Educational and psychological measurement. 55: 38-46.

Johnstone, J.N. 1981. Indicators of education systems. London: UNESCO.

JUAA. 2001. Japanese University Accreditation Association. [online] Available from: www.juaa.or.jp/english [2002, August 21]

Lind, D.A. and Mason, R.D. 1997. Basic statistics for business and economics. Chicago: McGraw-Hill.

Lim, D. 2001. Quality assurance in higher education: A study of developing countries. Burlington: Ashgate Publishing Company.

Lord, F.M. and Novick, M.R. 1968. Statistical theories of metal test scores. Massachusetts : Addison-Wesley Publishing company, Inc.

MacCallum, R.C. and et al. 1997. Studying multivariate change using multilevel models and latent curve models. Multivariate behavioral research. 32(3) : 215-253.

Massy, W. 1997. Teaching and learning quality-process review: The Hong Kong programme. Quality in higher education. 3(3): 249-262.

Meredith, W. and Tisak, J. 1990. Latent curve analysis. Psychometrika. 55(1) : 107-122.

Ministry of University Affairs (MUA). 2001. QA@MUA.THAILAND. Bangkok: Lardproa Press.

Ministry of University Affairs (MUA). 1998. Quality Assurance in Thailand. [online] Available from: www.qa.mua.go.th [2002, August 21].

Mok, K.H. 2000. Impact of globalization: A study of quality assurance systems of higher education in Hong Kong and Singapore. Comparative Education Review. 44(2): 148-174.

Neter, J. and Wasserman, W. 1974. Applied linear statistical models. Illinois : Richard D. Irwin Inc.

Newton, J. 1999. An evaluation of the impact of external quality monitoring on a higher education college (1993-98). Assessment & evaluation in higher education. 24(2): 215-235.

NZQA. 2002a. QA standard one. [online] Available from: www.nzqa.govt.nz [2002, August 21].

NZQA. 2002b. New Zealand Qualifications Authority. [online] Available from: www.nzqa.govt.nz [2002, August 21].

Office of the National Education Commission (ONEC). 2000. Education in Thailand 2000/2001. Bangkok: Kurusapa Lardprao Press.

Office of the National Education Commission (ONEC). 1999. National education act of B.E. 2542 (1999). Bangkok: Seven Printing Group.

Piascik, C. 1994. Applied calculus for business and the social and natural sciences. Info Access & Distribution Pte Ltd.

QAA. 2002. Quality Assurance Agency. [online] Available from: www.qaa.ac.uk [2002, August 21].

Quality Assurance Academic Affairs Department, TU. 2002: Quality assurance [online] Available from: www.tu.ac.th) [2002, August 21].

Raudenbush, S., Bryk, A. and Congdon, R. 2000. HLM 5 (Hierarchical linear and nonlinear Model. www.ssicentral.com.

Raudenbush, S. 1995. Hierarchical linear models to study the effects of social context on Development, edit by John Mordechai Gottman. The analysis of change. New Jersey: Lawrence Erlbaum Associates, Publishers: 165-201.

Rakov, T. 1993. Structural equation model for measuring residualized change and discerning patterns of growth or decline. Applied Psychological Measurement. 17 : 53-71.

Rogosa, D. and Willett, J.B. 1985. Understanding correlates of change by modeling individual differences in growth. Psychometrika. 50(2): 203-228.

Rogosa, D. and Saner, H. 1995. Longitudinal data analysis examples with random coefficient Models. Journal of educational and behavioral statistics. 20(2): 149-170.

Rogosa, D. 1995. Myths and methods : “Myths about longitudinal research” plus supplemental quesions, edit by John Mordechai Gottman. The analysis of change. New Jersey: Lawrence Erlbaum Associates, Publishers: 165-201.

Shin, J., Deno, S.L. and Espin, C. 2000. Technical adequacy of the maze task for curriculum-based measurement of reading growth. The journal of special education. 34(3): 164-172.

Silverstein, M. and Long, J.D. 1998. Trajectories of grandparents’ perceived solidarity with adult Grandchildren: A growth curve analysis over 23 years. Journal of marriage and the Family. 60: 912-923.

Soilman, I and Soliman, H. 1997. Academic workload and quality. Assessment & Evaluation in Higher Education. 22(2): 135-158.

SU. 2002. Saint John’s University background. [online] Available from: www.stjohn.ac.th [2002, August 21].

TU. 2002. Thammasat University history. [online] Available from: www.tu.ac.th [2002, August 21].

University Grants Committee of Hong Kong (UGC). 1996. Higher education in Hong Kong. Hong Kong: Government Printer.

Willamson, G.L., Appelbaum, M. and Epanchin, A. 1991. Longitudinal analyses of academic Achievement. Journal of educational measurement. 28(1): 61-76.

Willett, J.B. and Sayer, A.G. 1996. Cross-Domain analyses of change over time: Combining growth modeling and covariance structure analysis. edit by Marcoulides, G and Schumacker, R.E. Advanced structural equation modeling : issue and techniques. New Jersey : Lawrence Erlbaum Associates, Inc.: 125-157.

Willett, J.B. 1990. Measuring change : The difference score and beyond. edit by Walberg, H.J. and Haertel, G.D. The international encyclopedia of educational evaluation. Oxford : Pergamon Press: 632-637.

Wise, A.E. 1996. Building a system of quality assurance for the teaching profession: moving into the 21st century. Phi Delta Kappan. 78: 190-192.

Woodhouse, D. 1996. Quality assurance: International trends, preoccupations and features. Assessment & Evaluation in Higher Education. 21(4): 347-357.

Appendices

Appendix A

Research Instruments

Instrument 1

An Interview form for selecting major tasks of the instructor

Questions

1. “University Instructors’ Performance” are divided into 6 tasks: (1) Teaching, (2) Student advice, (3) Research and publications, (4) Academic service to community, (5) Preservation of arts and culture, and (6) Administration and academic self-improvement. Do you agree with this?

☐ Agree ☐ Disagree

Suggestions……

……

2. Subtasks in “Teaching” task, what subtasks should be combined for measuring instructors’ performance, which should be removed and what new subtasks should be added and how the definition and wording of the subtasks could be improved?

……

……

……

……

3. Subtasks in “Student advice” task, what subtasks should be combined for measuring instructors’ performance, which should be removed and what new subtasks should be added and how the definition and wording of the subtasks could be improved?

……

……

……

……

4. Subtasks in “Research and academic publications” task, what subtasks should be combined for measuring instructors’ performance, which should be removed and what new subtasks should be added and how the definition and wording of the subtasks could be improved?

……

……

……

……

5. Subtasks in “Academic service to community” task, what subtasks should be combined for measuring instructors’ performance, which should be removed and what new subtasks should be added and how the definition and wording of the subtasks could be improved?

……
……
……
……

6. Subtasks in “Preservation of arts and culture” task, what subtasks should be combined for measuring instructors’ performance, which should be removed and what new subtasks should be added and how the definition and wording of the subtasks could be improved?

……
……
……
……

7. Subtasks in “Administration and academic self-improvement” task, what subtasks should be combined for measuring instructors’ performance, which should be removed and what new subtasks should be added and how the definition and wording of the subtasks could be improved?

……
……
……
……

8. Suggestions

……
……
……
……

A list of subtasks for university instructors

Task	Subtask
Task 1: Teaching	
1.1 Teaching preparation	1.1 Has knowledge of assigned subjects
	1.2 Keeps up to date with assigned subjects
	1.3 Course outlines produced
	1.4 Lesson plans produced
	1.5 Gives course outlines to the students in the first day of the course
	1.6 Course outline consist of the objectives and guidelines for study
	1.7 Any notes, etc. hand out were pitched at the appropriate education level and length of learning period
	1.8 Instructional media prepared in advance
1.2 Teaching	1.9 Number of teaching hours per week
	1.10 Teaching methods are appropriate to the course's objectives, contents and students
	1.11 Teaching methods are suitable given the backgrounds and experiences of the students and the emphasis is on student learning
	1.12 Learning experiences were provided through many different methods
	1.13 Students were introduced to relevant sources and reference books
	1.14 Students were given opportunities to present their studies
	1.15 Keeps up to date with assigned subjects regularly
	1.16 Provides activities and environment that stimulate students' learning
	1.17 Appropriate use of instructional media
	1.18 On time for both the beginning and ending of teaching periods
1.3 Student assessment	1.19 Criteria for student assessment were decided in advance
	1.20 Criteria for student assessment were announced to students
	1.21 Use many different assessment methods
	1.22 Examination administered on time
	1.23 Developing tests and examinations
	1.24 Grades reported to faculty on time
1.4 Teaching evaluation	1.25 Results of evaluation by students
	1.26 Self-evaluation in teaching
	1.27 Teaching evaluation by peer, senior staff and students
	1.28 Use of evaluation results
Task 2: Student Advice	2.1 Number of advice hours per week
	2.2 Number of advisees
	2.3 Give counseling or advice in academic, career and personal issues
	2.4 Specific times set aside for advising students

Task	Subtask
Task 3: Research and Academic Publications	
	3.1 Number of published research papers
	3.2 the amount of research funds held
	3.3 the amount of research funds received from outside
	3.4 Number of rewards
	3.5 the amount of citations
	3.6 Number of text written
	3.7 Number of academic books written
	3.8 Number of academic articles published in academic journals
	3.9 Number of academic articles presented in academic conference
	3.10 Classification for academic publications
Task 4: Academic Service to Community	4.1 Number of invited lecturers (in or out university)
	4.2 Number of times serve outside the university
	4.3 Number of service hours per week
	4.4 Academic service to community on areas of specialization
	4.5 Academic service to community fits into faculty plan or policy
	4.6 Academic service to community was allowed from faculty
	4.7 Satisfaction of the persons who received service
Task 5: Preservation of Arts and Culture	5.1 Establishes morals, ethics, and culture into students
	5.2 Introduces or inserts morals, ethics and culture in teaching and learning methodology
	5.3 Be good role models for students
	5.4 Participant constantly in arts and culture preservation activities
	5.5 Participates and support art and cultural activities
	5.6 Organizes projects that help preserve arts and culture occasionally
	6.1 Number of times attended at meetings
Task 6: Administration and Academic Self-Improvement	6.2 Attendance at department/faculty/ university meetings
	6.3 Number of permanent committee memberships
	6.4 Number of temporary committee memberships
	6.5 Keeps systematic documentary evidence of work undertaken
	6.6 Number of times attended academic training/seminars/conferences
	6.7 Training/seminars/conferences relevant to duties
	6.8 Reports and disseminates to other instructors in their department/faculty
	6.9 Implementation of the knowledge gained from training/seminars/conferences

Instrument 2

An Interview form for confirming the subtasks identified

Directions: Please give your opinions on how each subtask be appropriate for measuring instructor's performance by ticking ✓ in "Opinion" column in the table below:

5 = Strongly agree – Subtask is very appropriate for measuring instructors' performance

4 = Agree – Subtask is appropriate for measuring instructors' performance

3 = Not sure – No idea of appropriateness of subtask for measuring instructors' performance

2 = Disagree – Subtask is not appropriate for measuring instructors' performance

1 = Strongly disagree – Subtask is not at all appropriate for measuring instructors' performance

Subtask	Opinion				
	5	4	3	2	1
Task 1: Teaching					
1.1 Course outlines produced					
1.2 Lesson plans produced					
1.3 Teaching load (number of teaching hours per week)					
1.4 Use of appropriate teaching method for students and subjects					
1.5 Provides activities and environment that stimulate students' learning					
1.6 Appropriate use of instructional media					
1.7 Number of texts used					
1.8 Number of assessment methods used					
1.9 Examination administered on time					
1.10 Grades reported to faculty on time					
1.11 Results of evaluation by students					
1.12 Use of evaluation results					
Task 2: Student Advice					
2.1 Specific times set aside for advising students					
2.2 Depth of advice offered					
2.3 Number of advice hours per week					
2.4 Number of advisees					
Task 3: Research and Academic Publications					
3.1 Number of academic articles published or presented in academic conferences/seminars					
3.2 Number of research projects contributed to					
3.3 Number of text or other academic books written					
3.4 Number of teaching materials written					
3.5 Quality of academic work					
3.6 Status of publications					

Subtask	Opinion				
	5	4	3	2	1
Task 4: Academic Service to Community					
4.1 Number of invited lecturers (in or out university)					
4.2 Number of service hours per week					
4.3 Academic service to community on areas of specialization					
4.4 Academic service to community fits into faculty plan or policy					
Task 5: Preservation of Arts and Culture					
5.1 Participant in arts and culture preservation activities					
5.2 Establishes morals, ethics, and culture into students					
5.3 Contribution to arts and culture preservation activities / projects for the community					
Task 6: Administration and Academic Self-Improvement					
6.1 Number of permanent committee memberships					
6.2 Number of temporary committee memberships					
6.3 Attendance at department/faculty/ university meetings					
6.4 Keeps systematic documentary evidence of work undertaken					
6.5 Number of times attended academic training/seminars/conferences					
6.6 Implementation of the knowledge gained from training/seminars/conferences					

Suggestions:

……

……

……

……

……

……

……

……

……

……

Thank you very much in deep

Jitlekha Teerajarmorn

Instrument 3

A questionnaire to identify weighting of tasks and subtasks

For thesis namely

A DEVELOPMENT OF MEASURING METHODS OF UNIVERSITY INSTRUCTOR'S PERFORMANCE-CHANGE SCORES DUE TO THE IMPLEMENTATION OF THE QUALITY ASSURANCE: A CASE STUDY OF 2 UNIVERSITIES.

Directions:

1. The objective of this questionnaire is to identify the weighting of tasks and subtasks that indicate university instructors' performance.

2. This questionnaire is compose of 2 parts:

Part 1 The weighting of university instructor's tasks, the tasks are 1) Teaching, 2) Student advice, 3) Research and publications, 4) Academic service to community, 5) Preservation of arts and culture, and 6) Administration and academic self-improvement

Part 2 The weighting of subtasks (35 subtasks)

3. Please complete the questions:

Part 1 Please identify the weighting of each task on the assumption that the total score of six tasks is 100.

Part 2 Please identify the weighting of subtasks in a task on the assumption that the total score of a task is 100.

****

Please answer every question
and please return before 25 September 2001

****

Thesis advisor: Prof. Utumporn Jamornmann
Researcher : Miss Jitlekha Teerajarmorn (Tel. (02)9422093, (02)2182586)
(Doctoral degree student in department of educational research, Chulalongkorn University)

Part 1: The weighting of instructors' tasks

Direction: Please identify the weighting of each task that the total score of six tasks is **100**.

Task	Weighting
1. Teaching	
2. Student advice	
3. Research and academic publications	
4. Academic service to community	
5. Preservation of arts and culture	
6. Administration and academic self-improvement	
Total	**100**

Part 1: The weighting of instructors' tasks

Direction: Please identify the weighting of subtasks in a task on the assumption that the total score of a task is **100**.

Task 1: Teaching

Subtask	Weighting
1.1 Course outlines produced	
1.2 Lesson plans produced	
1.3 Teaching load (number of teaching hours per week)	
1.4 Use of appropriate teaching method for students and subjects	
1.5 Provides activities and environment that stimulate students' learning	
1.6 Appropriate use of instructional media	
1.7 Number of texts used	
1.8 Number of assessment methods used	
1.9 Examination administered on time	
1.10 Grades reported to faculty on time	
1.11 Results of evaluation by students	
1.12 Use of evaluation results	
Total	**100**

Task 2: Student Advice

Subtask	Weighting
2.1 Specific times set aside for advising students	
2.2 Depth of advice offered	
2.3 Number of advice hours per week	
2.4 Number of advisees	
Total	**100**

Task 3: Research and Academic Publications

Subtask	Weighting
3.1 Number of academic articles published or presented in academic conferences/seminars	
3.2 Number of research projects contributed to	
3.3 Number of text or other academic books written	
3.4 Number of teaching materials written	
3.5 Quality of academic work	
3.6 Status of publications	
Total	**100**

Task 4: Academic Service to Community

Subtask	Weighting
4.1 Number of invited lecturers (in or out university)	
4.2 Number of service hours per week	
4.3 Academic service to community on areas of specialization	
4.4 Academic service to community fits into faculty plan or policy	
Total	**100**

Task 5: Preservation of Arts and Culture

Subtask	Weighting
5.1 Participant in arts and culture preservation activities	
5.2 Establishes morals, ethics, and culture into students	
5.3 Contribution to arts and culture preservation activities / projects for the community	
Total	**100**

<u>Task 6</u>: Administration and Academic Self-improvement

Subtask	Weighting
6.1 Number of permanent committee memberships	
6.2 Number of temporary committee memberships	
6.3 Attendance at department/faculty/ university meetings	
6.4 Keeps systematic documentary evidence of work undertaken	
6.5 Number of times attended academic training/seminars/conferences	
6.6 Implementation of the knowledge gained from training/seminars/conferences	
Total	**100**

Thank you very much for your kindness

Instrument 4

A form for collecting data from documents

Background

1. Sex (1) Male (2) Female
2. Academic post (1) Lecturer (2) Assistant Prof. (3) Associate Prof. (4) Professor
3. Degrees (1) Bachelor (2) Master (3) Doctor
4. Teaching experience in higher education ……. years
5. Faculty …… Department ……..

Data on Subtasks

Teaching	Academic year 1998						Academic year 1999						Academic year 2000					
- Number of assigned subjects																		
- Number of course outline produced																		
Subject →	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
- Teach hours at undergraduate (theory)																		
- Teach hours at graduate (theory)																		
- Teach hours at undergraduate (practice)																		
- Teach hours at graduate (practice)																		
- Number of texts used																		
- Number of assessment methods used																		
- Results of evaluation by students																		
Student advice	Academic year 1998						Academic year 1999						Academic year 2000					
- Number of advisees																		
Research and academic publications	Academic year 1998						Academic year 1999						Academic year 2000					
- Number of academic articles																		
- Number of research projects contributed to																		
- Number of text or books written																		
- Number of teaching materials written																		
Academic service to community	Academic year 1998						Academic year 1999						Academic year 2000					
- Number of invited lecturers																		
Administration and self-improvement	Academic year 1998						Academic year 1999						Academic year 2000					
- Number of permanent committee memberships																		
- Number of temporary committee memberships																		
- Number of times attended academic training/seminars/conferences																		

Instrument 5

A self-report questionnaire

For thesis namely

A DEVELOPMENT OF MEASURING METHODS OF UNIVERSITY INSTRUCTOR'S PERFORMANCE-CHANGE SCORES DUE TO THE IMPLEMENTATION OF THE QUALITY ASSURANCE: A CASE STUDY OF 2 UNIVERSITIES.

Directions:

1. This questionnaire on 6 tasks on university instructors' performance: 1) Teaching, 2) Student advice, 3) Research and publications, 4) Academic service to community, 5) Preservation of arts and culture, and 6) Administration and academic self-improvement consisted of 21 items.

2. Each item has 5 choices: ①, ②, ③, ④ and ⑤. In each academic year, please select **only one choice** in each item, which you think it showed exactly your performance in each academic year by ticking ✓ in "Academic year" column in the table below:

Example: Suppose Instructor C produced his lesson plans as follows:

- Academic year 1998, Taught 3 subjects and produced lesson plan 1 subject
- Academic year 1999, Taught 4 subjects and produced lesson plan 2 subjects
- Academic year 2000, Taught 4 subjects and produced lesson plan 4 subjects

	Academic year		
1. Lesson plans produced	1998	1999	2000
① None (0%)			
② Some (1-33%)	✓		
③ Many (34-66%)		✓	
④ Almost all (67-99%)			
⑤ All (100%)			✓

****

Please answer every question
and please return before 21 September 2001

****

Thesis advisor: Prof. Utumporn Jamornmann

Researcher : Miss Jitlekha Teerajarmorn (Tel. (02)9422093, (02)2182586)

(Doctoral degree student in department of educational research, Chulalongkorn University)

Task 1: Teaching	Academic year		
1. Lesson plans produced	1998	1999	2000
① None (0%)			
② Some (1-33%)			
③ Many (34-66%)			
④ Almost all (67-99%)			
⑤ All (100%)			
2. Use of appropriate teaching method for students and subjects	1998	1999	2000
① Only use one method for every subjects			
② Use different methods sometimes			
③ Use many difference methods			
④ Use many difference methods that emphasize on students' learning			
⑤ Use many difference methods that emphasize on students' learning, and use innovation technologies			
3. Provides activities and environment that stimulate students' learning such as demonstration, role-play, practice etc.	1998	1999	2000
① None (0%)			
② Some (1-33%)			
③ Many (34-66%)			
④ Almost all (67-99%)			
⑤ All (100%)			
4. Appropriate use of instructional media such as pictographs, real objects, models, video, charts, internet etc.	1998	1999	2000
① None (0%)			
② Some (1-33%)			
③ Many (34-66%)			
④ Almost all (67-99%)			
⑤ All (100%)			
5. Examination administered on time	1998	1999	2000
① None (0%)			
② Some (1-33%)			
③ Many (34-66%)			
④ Almost all (67-99%)			
⑤ All (100%)			
6. Grades reported to faculty on time	1998	1999	2000
① None (0%)			
② Some (1-33%)			
③ Many (34-66%)			
④ Almost all (67-99%)			
⑤ All (100%)			

Task 1: Teaching (continue)	Academic year		
7. Use of evaluation results	1998	1999	2000
① No evaluation			
② Only interested in self evaluation			
③ Allows / organizes evaluation by others (e.g. students, peers, heads)			
④ Interested in evaluation results			
⑤ Use evaluation results to improve teaching			
Task 2: Student Advice	Academic year		
8. Specific times set aside for advising students	1998	1999	2000
① No specific time set			
② Have specific time set but never stay			
③ Have specific time set and sometimes stay			
④ Have specific time set and always stay			
⑤ Have specific time set, always stay, and have extra time if students need			
9. Depth of advice offered	1998	1999	2000
① No time			
② Only in the subjects			
③ In the subjects and other matters relating teaching/learning such as registration, scholarship etc.			
④ All matters including personal			
⑤ Advice, feed back and contact other offices			
10. Number of advice hours per week	1998	1999	2000
① None			
② 0.01 – 2 hours			
③ 2.01 – 4 hours			
④ 4.01 – 6 hours			
⑤ More than 6 hours			
Task 3: Research and Academic Publications (covering conducting research and writing textbooks, other academic books, academic articles, teaching supplement materials, and other innovations)	Academic year		
11. Quality of academic work (majority)	1998	1999	2000
① No academic work			
② Collecting knowledge			
③ Including creative and new knowledge			
④ Including new knowledge accepted nationally			
⑤ Including new knowledge accepted internationally			

Task 3: Research and Academic Publications (continue)	Academic year		
12. Status of publications (majority)	1998	1999	2000
① No academic work			
② Exist work but no publish			
③ Present in academic conference			
④ Publish in journal / book at national level			
⑤ Publish in journal / book at international level			
Task 4: Academic service to community including giving invited lectures, being a visiting lecturer, disseminating knowledge to the wider community by mass communication, giving academic advice to the government or private organisations, and arranging academic activities such as professional meetings, in-service training, seminars, and exhibitions	Academic year		
13. Academic service to community on areas of specialization	1998	1999	2000
① None (0%)			
② Some (1-33%)			
③ Many (34-66%)			
④ Almost all (67-99%)			
⑤ All (100%)			
14. Academic service to community fits into faculty plan or policy	1998	1999	2000
① None (0%)			
② Some (1-33%)			
③ Many (34-66%)			
④ Almost all (67-99%)			
⑤ All (100%)			
15. Number of service hours per week (not include invited lecturer)	1998	1999	2000
① None			
② 0.01 – 3 hours			
③ 3.01 – 6 hours			
④ 6.01 – 9 hours			
⑤ More than 9 hours			
Task 5: Preservation of Arts and Culture	Academic year		
16. Participant on arts and culture preservation activities such as present food to a Buddhist priest, dress Thai fabric, participate Thai festivals etc.	1998	1999	2000
① None (0%)			
② Some (1-33%)			
③ Many (34-66%)			
④ Almost all (67-99%)			
⑤ All (100%)			

Task 5: Preservation of Arts and Culture (continue)	Academic year		
17. Establishes moral ethics and culture into students	1998	1999	2000
① Never do			
② Insert in teaching sometimes			
③ Insert in teaching as possible			
④ Insert in teaching and warn students when they do the wrong things			
⑤ Insert in teaching, warn students, and arrange activities that established moral, ethics, and culture into them			
18. Contribution to art and culture preservation activities/projects for the community	1998	1999	2000
① Never do			
② Cooperating in the activities/projects			
③ Being membership of activities/projects			
④ Being leader of activities/projects			
⑤ Persuading other persons to join activities/projects			
Task 6: Administration and Academic Self-improvement	Academic year		
19. Attending at department/faculty/university meetings	1998	1999	2000
① Little (0-24%)			
② Some (25-49%)			
③ Many (50-74%)			
④ Always (75-99%)			
⑤ All (100%)			
20. Keeps systematic documentary evidence of work undertaken	1998	1999	2000
① Never keep			
② Keep some but no system			
③ Keep many but not all in system			
④ Keep many and all in system			
⑤ Keep all and all in system			
21. Implementation of the knowledge gained from training/seminars/ conferences	1998	1999	2000
① Never			
② Report to faculty			
③ Transfer knowledge to peers			
④ Utilize some knowledge			
⑤ Utilize knowledge that useful for teaching and research			

Thank you very much for your kindness

Instrument 6

Example

An interview form for validating performance scores and performance change scores

For Head of department T94

1. Instructor's performance score and performance change scores

Instructor	Performance score (Total 400)			Difference score		Change score	
	Academic year 1998	Academic year 1999	Academic year 2000	42 - 41	43 – 42	Method 1 (Max 150)	Method 2 (Max 200)
Instructor9031	150	148	182	- 2	34	13.9	15.5
Instructor9032	193	218	223	25	5	14.9	14.2
Instructor9035	218	239	244	21	5	13.7	12.4
Faculty	178	196	207	18	11	13.8	13.9

1. In your opinion, the performance scores of each instructor in your department (shown in above table) are valid or not valid?

Instructor9031 ❒ Valid ❒ Not valid

Instructor9032 ❒ Valid ❒ Not valid

Instructor9035 ❒ Valid ❒ Not valid

2. Please compare change scores obtained from Method 1 and Method 2, what method was more validate for calculating change score of each instructor during academic year 1998-2000?

Decision	Instructor9031	Instructor9032	Instructor9033
Method 1 <u>better than</u> Method 2			
Method 1 <u>as same as</u> Method 2			
Method 1 <u>less than</u> Method 2			

Thank you very much for your kindness

Jitlekha Teerajarmorn

P.S. In the real interview form, the researcher used the name of instructors

Instrument 7

Example

An interview form for validating performance change scores

For Instructor 9032

1. Instructor's performance score and performance change scores

Instructor	Performance score (Total 400)			Difference score		Change score	
	Academic year 1998	Academic year 1999	Academic year 2000	42 - 41	43 – 42	Method 1 (Max 150)	Method 2 (Max 200)
Instructor9032	193	218	223	25	5	14.9	14.2
Instructor A	225	248	255	23	7	16.1	13.9
Faculty	178	196	207	18	11	13.8	13.9

Question: Please compare change scores obtained from Method 1 and Method 2, what method was more validate for calculating your change score during academic year 1998-2000?

- ❒ Method 1 better than Method 2
- ❒ Method 1 as same as Method 2
- ❒ Method 1 less than Method 2

Thank you very much for your kindness

Jitlekha Teerajarmorn

Appendix B

Senior and Experienced Staff Who Gave Data

Appendix B
Senior and Experienced Staff Who Gave Data

B.1 Senior Staff and Experienced Instructors (Subgroup 1.1)

มหาวิทยาลัยธรรมศาสตร์

1. ผศ.ดร.วีรยา ฉิมอ้อย	รองคณบดีฝ่ายวิชาการ คณะวิศวกรรมศาสตร์
2. ผศ.ดร.ปรีชา วาณิชยเศรษฐกุล	หัวหน้าสถานวิทยาศาสตร์พรีคลินิก, หัวหน้าโครงการประกันคุณภาพ คณะแพทยศาสตร์
3. คุณสุภาพ ดวงไสว	ผู้อำนวยการกองการบริการการศึกษา, รับผิดชอบด้านการประกันคุณภาพระดับมหาวิทยาลัย
4. ผศ.ดร.สุกัญญา บำรุงสุข	รองคณบดีฝ่ายวิชาการ คณะศิลปศาสตร์
5. ผศ.ดร.อุดม รัฐอมฤต	รองคณบดีฝ่ายบริหาร คณะนิติศาสตร์
6. ผศ.ดร.ฉวีวรรณ ประจวบเหมาะ	รองคณบดีฝ่ายวิชาการ คณะสังคมวิทยาและมานุษยวิทยา
7. รศ.มาลี บุญศิริพันธ์	อาจารย์ประจำคณะวารสารศาสตร์และสื่อสารมวลชน

มหาวิทยาลัยเซนต์จอห์น

1. ดร.สุวิชากร ชินะผา	รองอธิการบดีฝ่ายวิชาการ
2. อ.อัญชลี เหลืองอ่อน	ผู้อำนวยการสำนักประกันคุณภาพและข้อมูล
3. อ.ธนีนาฏ ณ สุนทร	อดีตคณบดีคณะศิลปศาสตร์ / ประธานหลักสูตรศึกษาศาสตร์มหาบัณฑิต
4. อ.ธิดา อนันตเสรี	รองคณบดีฝ่ายบริหาร คณะบริหารธุรกิจ
5. อ.เนาวรัตน์ เทพอาสน์	รองคณบดีฝ่ายวิชาการ คณะนิเทศศาสตร์
6. อ.ปราการ วยาจุต	ผู้อำนวยการสำนักบริหารงานบุคคล
7. อ.สุมนา ประคองวิทยา	อาจารย์ประจำคณะวิศวกรรมศาสตร์ / คณะกรรมการกลุ่มดำเนินงานประกันคุณภาพการศึกษา
8. อ.บรรพต ปานจันทร์	อาจารย์ประจำ / เลขานุการศูนย์ศิลปวัฒนธรรม

B.2 Senior Staff and Experienced Instructors (Subgroup 1.2)

มหาวิทยาลัยธรรมศาสตร์

1. รศ.กิติมา สุรสนธิ	รองคณบดีฝ่ายวิชาการ คณะวารสารศาสตร์และสื่อสารมวลชน
2. ผศ.ดร.วีรยา ฉิมอ้อย	รองคณบดีฝ่ายวิชาการ คณะวิศวกรรมศาสตร์
3. ผศ.ดร.ฉวีวรรณ ประจวบเหมาะ	รองคณบดีฝ่ายวิชาการ คณะสังคมวิทยาและมานุษยวิทยา
4. ผศ.ดร.ทองดี เล็กโสภี	รองคณบดีฝ่ายวิชาการ คณะวิทยาศาสตร์
5. รศ.นรีทิพทย์ ทุ่งกาวี	อาจารย์ประจำคณะพาณิชยศาสตร์และการบัญชี
6. อ.ดร.จิรรัตน์ ธีระวราพฤกษ์	อาจารย์ประจำคณะวิศวกรรมศาสตร์
7. อ.ดร.อนุชา ทีรคานนท์	อาจารย์ประจำคณะวารสารศาสตร์และสื่อสารมวลชน
8. อ.ดร.เพียงจันทร์ เศวตศรีสกุล	อาจารย์ประจำคณะแพทยศาสตร์

มหาวิทยาลัยเซนต์จอห์น

1. อ.อัญชลี เหลืองอ่อน	ผู้อำนวยการสำนักประกันคุณภาพและข้อมูล
2. อ.ธนีนาฏ ณ สุนทร	อดีตคณบดีคณะศิลปศาสตร์ / ประธานหลักสูตรศึกษาศาสตร์มหาบัณฑิต
3. อ.ธิดา อนันตเสรี	รองคณบดีฝ่ายบริหาร คณะบริหารธุรกิจ
4. อ.เนาวรัตน์ เทพอาสน์	รองคณบดีฝ่ายวิชาการ คณะนิเทศศาสตร์
5. อ.ปราการ วยาจุต	ผู้อำนวยการสำนักบริหารงานบุคคล
6. ดร.หอม คลายานนท์	ผู้อำนวยการสำนักวิจัย / อดีตคณบดีคณะศิลปศาสตร์
7. อ.วิลาสินี โฆษจันทร	ผู้อำนวยการศูนย์ศิลปวัฒนธรรม / อาจารย์ประจำคณะนิเทศศาสตร์
8. อ.พรหมอาสน์ ประดิษฐ์	ผู้อำนวยการบัณฑิตศึกษา
9. อ. ธรรมรงค์ สุขชื่น	ประธานคณะกรรมการกลุ่มดำเนินงานประกันคุณภาพ

B.3 Experts for subtask weighting (Group 2)

มหาวิทยาลัยธรรมศาสตร์

1. รศ.ดร.มัทนา พนานิรามัย	รองคณบดีฝ่ายวิชาการ คณะเศรษฐศาสตร์
2. ผศ.ดร.กิติพัฒน์ นนทปัทมะดุลย์	รองคณบดีฝ่ายวิชาการและวิจัย คณะสังคมสงเคราะห์
3. ผศ.ดร.อุดม รัฐอมฤต	รองคณบดีฝ่ายบริหาร คณะนิติศาสตร์
4. รศ.กิติมา สุรสนธิ	รองคณบดีฝ่ายวิชาการ คณะวารสารศาสตร์และสื่อสารมวลชน
5. ผศ.ดร.วีรยา ฉิมอ้อย	รองคณบดีฝ่ายวิชาการ คณะวิศวกรรมศาสตร์
6. ผศ.ดร.ฉวีวรรณ ประจวบเหมาะ	รองคณบดีฝ่ายวิชาการ คณะสังคมวิทยาและมานุษยวิทยา
7. ผศ.ดร.ทองดี เล็กโสภี	รองคณบดีฝ่ายวิชาการ คณะวิทยาศาสตร์
8. ผศ.ดร.ยุพิน ส่งไพศาล	รองคณบดีฝ่ายวิชาการ คณะทันตแพทยศาสตร์
9. อ.นพ.พีระพงศ์ กิติภาวงค์	รองคณบดีฝ่ายวิชาการ คณะแพทยศาสตร์

มหาวิทยาลัยเซนต์จอห์น

1. ดร.สุวิชากร ชินะผา	รองอธิการบดีฝ่ายวิชาการ
2. อ.อัญชลี เหลืองอ่อน	ผู้อำนวยการสำนักประกันคุณภาพและข้อมูล
3. อ.ธนีนาฏ ณ สุนทร	อดีตคณบดีคณะศิลปศาสตร์ / ประธานหลักสูตรศึกษาศาสตร์มหาบัณฑิต
4. อ.ธิดา อนันตเสรี	รองคณบดีฝ่ายบริหาร คณะบริหารธุรกิจ
5. อ.เนาวรัตน์ เทพอาสน์	รองคณบดีฝ่ายวิชาการ คณะนิเทศศาสตร์
6. อ.ปราการ วยาจุต	ผู้อำนวยการสำนักบริหารงานบุคคล
7. ดร.หอม คลายานนท์	ผู้อำนวยการสำนักวิจัย / อดีตคณบดีคณะศิลปศาสตร์
8. อ.วิลาสินี โฆษจันทร	ผู้อำนวยการศูนย์ศิลปวัฒนธรรม / อาจารย์ประจำคณะนิเทศศาสตร์

B.4 Experienced Instructors for improving the self-report questionnaire

1. อ.อัญชลี เหลืองอ่อน	ผู้อำนวยการสำนักประกันคุณภาพและข้อมูล มหาวิทยาลัยเซนต์จอห์น
2. อ.ธนีนาฏ ณ สุนทร	อดีตคณบดีคณะศิลปศาสตร์ / ประธานหลักสูตรศึกษาศาสตร์มหาบัณฑิต มหาวิทยาลัยเซนต์จอห์น
3. อ.ธิดา อนันตเสรี	รองคณบดีฝ่ายบริหาร คณะบริหารธุรกิจ มหาวิทยาลัยเซนต์จอห์น
4. อ.เนาวรัตน์ เทพอาสน์	รองคณบดีฝ่ายวิชาการ คณะนิเทศศาสตร์ มหาวิทยาลัยเซนต์จอห์น
5. รศ.กิติมา สุรสนธิ	รองคณบดีฝ่ายวิชาการ คณะวารสารศาสตร์และสื่อสารมวลชน มหาวิทยาลัยธรรมศาสตร์
6. อ.ดร.จิรรัตน์ ธีระวราพฤกษ์	อาจารย์ประจำคณะวิศวกรรมศาสตร์
7. รศ.ดร.จงกล เที่ยงดาห์	อาจารย์ประจำคณะเภสัชศาสตร์ มหาวิทยาลัยมหิดล
8. อ.ดร.เพ็ญศรี ธีระวราพฤกษ์	ที่ปรึกษา SEAMEO Regional Center for Higher Education and Development / อดีตอาจารย์มหาวิทยาลัย (ประสบการณ์ทำงานมากกว่า 30 ปี)

Appendix C

Criteria for Scoring Subtasks

Appendix C
Criteria for Scoring Subtasks

University instructor's performance is divided into 6 tasks and 35 subtasks. The criteria for scoring subtasks are as follows:

Task 1: Teaching

Teaching is comprised of preparation, teaching, student assessment, and teaching evaluation. The task is composed of 12 subtasks as below:

1.1) Course outlines produced: The percentage of assigned subjects that the instructor produced course outlines in an academic year.

Course outline produced	Score
None (0%)	0
Some (1-33%)	1
Many (34-66%)	2
Almost all (67-99%)	3
All (100%)	4

1.2) Lesson plans produced: The percentage of assigned subjects that the instructor produced lesson plans in an academic year.

Lesson plan produced	Score
None (0%)	0
Some (1-33%)	1
Many (34-66%)	2
Almost all (67-99%)	3
All (100%)	4

1.3) Teaching load: The number of teaching hours per week (H) in an academic year that calculated from $H = A + (1/2)B + (3/2)C + (3/4)D$

Where A = number of teaching hours per week on theoretical subjects at undergraduate level.
B = number of teaching hours per week on practical subjects at undergraduate level.
C = number of teaching hours per week on theoretical subjects at graduate level.
D = number of teaching hours per week on practical subjects at graduate level.

Teaching Load	Score
Not over 3 hours	0
3.01 – 6 hours	1
6.01 – 9 hours	2
9.01 – 12 hours	3
More than 12 hours	4

1.4) Use of appropriate teaching method for students and subjects: The instructor used many different methods that emphasize on students' learning and also use innovation technologies.

Use of appropriate teaching method	Score
Only use one method for every subjects	0
Use different methods sometimes	1
Use many difference methods	2
Use many difference methods that emphasize on students' learning	3
Use many difference methods that emphasize on student's learning, and use innovation technologies	4

1.5) Provides activities and environment that stimulate students' learning: Frequency that the instructor provided activities and environment such as demonstration, role-play, discussion, try out, virtual model, practice, etc. that were appropriate for stimulating students' learning.

Provides activities and environment	Score
None (0%)	0
Some (1-33%)	1
Many (34-66%)	2
Almost all (67-99%)	3
All (100%)	4

1.6) Appropriate use of instructional media: Frequency that the instructor used instructional media appropriately to stimulate students' learning such as pictographs, real objects, models, pictures, charts, video, internet, etc.

Appropriate use of instructional media	Score
None (0%)	0
Some (1-33%)	1
Many (34-66%)	2
Almost all (67-99%)	3
All (100%)	4

1.7) Number of texts used: Average of number of texts that the instructor used in each subject.

Number of texts used	Score
1 – 3	0
4 - 6	1
7 - 9	2
10 - 12	3
More than 12	4

1.8) Number of assessment methods used: Average of number of assessment methods such as mid-term examination, final examination, class participation, reports, exercises, test, class attendance, etc. that the instructor used in each subject.

Number of assessment methods used	Score
Less than 3 methods	0
3 methods	1
4 methods	2
5 methods	3
More than 5 methods	4

1.9) Examination administered on time: The percentage of assigned subjects that the instructor administered examinations on time in an academic year.

Examination administered on time	Score
None (0%)	0
Some (1-33%)	1
Many (34-66%)	2
Almost all (67-99%)	3
All (100%)	4

1.10) Grades reported to faculty on time: The percentage of assigned subjects that instructor reported grades to faculty on time in an academic year.

Grades reported to faculty on time	Score
None (0%)	0
Some (1-33%)	1
Many (34-66%)	2
Almost all (67-99%)	3
All (100%)	4

1.11) Results of evaluation by students: Average of evaluation results by students in all subjects taught by the instructor in an academic year.

Results of evaluation by students		Score
Need improvement	(1.0 – 1.5)	0
Moderate	(1.6 – 2.5)	1
Good	(2.6 – 3.5)	2
Very good	(3.6 – 4.5)	3
Excellent	(4.6 – 5.0)	4

1.12) Use of evaluation results: The amount of evaluation that the instructor permitted and the use made of the results.

Use of evaluation results	Score
No evaluation	0
Only interested in self evaluation	1
Allows / organizes evaluation by others (e.g. students, peers, head of department)	2
Interested in evaluation results	3
Use evaluation results to improve teaching	4

Task 2: Student Advice

The task is composed of 4 subtasks as follows:

2.1) **Specific times set aside for advising students:** The instructor had specific time set aside for advising students and he stays in his office in that time.

Specific times set aside for advising students	Score
No specific time	0
Have specific time set but never stay	1
Have specific time set and sometimes stay	2
Have specific time set and always stay	3
Have specific time set, always stay, and have extra time if students need	4

2.2) **Depth of advice offered:** The depth of advice that instructor often offer to students.

Depth of advice offered	Score
No time	0
Only in the subjects	1
In the subjects and other matters relating teaching / learning such as registration, scholarship, etc.	2
All matters including personal	3
Advice, feed back and contact other offices	4

2.3) **Number of advice hours per week:** Average of advice hours per week that instructor advised students in academy and personal.

Number of advice hours per week	Score
None	0
0.01 – 2 hours	1
2.01 – 4 hours	2
4.01 – 6 hours	3
More than 6 hours	4

2.4) **Number of advisees:** Number of students under the instructor's advice depending on his faculty.

Social Science & Humanity	Law	Engineering & Science and Technology	Medicine, Nursing, Dentistry & Allied Health Science	Graduate School	Score
Less than 8	Less than 12	Less than 6	1 – 2	Less than 3	0
8 – 13	12 – 24	6 – 10	3 – 4	3	1
14 – 19	25 – 37	11 – 15	5 – 6	4	2
20 – 25	38 – 50	16 – 20	7 – 8	5	3
More than 25	More than 50	More than 20	More than 8	More than 5	4

Task 3: Research and Academic Publications

This task covers conducting research and writing textbooks, other academic books, academic articles, teaching supplement materials, and other innovations. The task is composed of 6 subtasks as follows:

3.1) Number of academic articles published or presented in academic conferences/ seminars: Number of academic articles that were published in journals or presented in national or international conference/seminars in an academic year.

Number of academic articles	Score
None	0
1	1
2	2
3	3
More than 3	4

3.2) Number of research projects contributed to: Number of research projects that the instructor contributed to in an academic year.

Number of research projects contributed to	Score
None	0
1	1
2	2
3	3
More than 3	4

3.3) Number of texts or other academic books written: Number of texts or other academic books that the instructor wrote in the first print in an academic year.

Number of texts or other academic books written	Score
None	0
1	1
2	2
3	3
More than 3	4

3.4) Number of teaching materials written: Number of teaching materials that the instructor wrote in the first print in an academic year.

Number of teaching materials written	Score
None	0
1	1
2	2
3	3
More than 3	4

3.5) Quality of academic work: The majority of the academic products being original, creative and innovative widely accepted nationally and internationally, etc.

Quality of academic work	Score
No academic work	0
Collecting knowledge	1
Including creative and new knowledge	2
Including new knowledge accepted nationally	3
Including new knowledge accepted internationally	4

3.6) Status of Publications: The majority of academic work spread out such as presenting in academic conference, publishing in journal / book at national or international level.

Status of Publications	Score
No academic work	0
Exist work but no publish	1
Present in academic conference	2
Publish in journal / book at national level	3
Publish in journal / book at international level	4

Task 4: Academic Service to Community

There are many ways an instructor can give academic service to the community including being invited lecturer, being a visiting lecturer, disseminating knowledge to the wider community by mass communication, giving academic advice to the government or private organizations, and arranging academic activities such as professional meetings, in-service training, seminars, and exhibitions. The task is composed of 4 subtasks as follows:

4.1) Number of times to be invited lecturers: Number of times that the instructor was invited to be lecturers inside or outside the university.

Number of times to be invited lecturers	Score
None	0
1 - 3	1
4 - 6	2
7 - 9	3
More than 9	4

4.2) Number of service hours per week: Number of hours per week that the instructor worked on academic service (not including invited lecturer).

Number of service hours per week	Score
None	0
0.01– 3 hours	1
3.01 – 6 hours	2
6.01 – 9 hours	3
More than 9 hours	4

4.3) Academic service to community on areas of specialization: The percentage of academic services, which the instructor serviced to community, related on areas of specialization.

Academic service to community on area of specialization	Score
None (0%)	0
Some (1-33%)	1
Many (34-66%)	2
Almost all (67-99%)	3
All (100%)	4

4.4) Academic service to community fits into faculty plan or policy: The percentage of academic services, which the instructor serviced to community, fitted into faculty plan or policy.

Academic service to community fits into faculty plan or policy	Score
None (0%)	0
Some (1-33%)	1
Many (34-66%)	2
Almost all (67-99%)	3
All (100%)	4

Task 5: Preservation of Arts and Culture

The task is composed of 3 subtasks as follows:

5.1) Participant in arts and culture preservation activities: Frequency that the instructor participated in arts and culture preservation activities such as present food to a Buddhist priest, dress Thai fabric, participate Thai festivals, present robes to Buddhist monks at a temple etc.

Participant in arts and culture preservation activities	Score
None (0%)	0
Some (1-33%)	1
Many (34-66%)	2
Almost all (67-99%)	3
All (100%)	4

5.2) Establishes morals, ethics, and culture into students: The level of instructor's performance for establishing moral, ethics, and culture into students.

Establishes morals, ethics, and culture into students	Score
Never do	0
Insert in teaching sometimes	1
Insert in teaching as possible	2
Insert in teaching and warn students when they do the wrong things	3
Insert in teaching, warn students, and arrange activities that established moral, ethics, and culture into them	4

5.3) Contribution to art and culture preservation activities/projects for the community: The level of instructor's contribution to activities/ projects about preservation of arts and culture for the community.

Contribution to activities/projects for the community	Score
Never do	0
Cooperating in the activities/projects	1
Being membership of activities/projects	2
Being leader of activities/projects	3
Persuading other persons to join activities/projects	4

Task 6: Administration and Academic Self-improvement

The task instructor work assigned to the instructor by the Faculty or university besides main task. The task is composed of 6 subtasks as follows:

6.1) Number of permanent committee memberships: Number of permanent committee memberships that the instructor worked on every week or month until the end of assignment.

Member of permanent committee memberships	Score
None	0
1	1
2	2
3	3
More than 3	4

6.2) Number of temporary committee memberships: Number of temporary committee memberships that the instructor work on a short period such as diploma presentation ceremony, sports committee etc.

Number of temporary committee memberships	Score
None	0
1	1
2	2
3	3
More than 3	4

6.3) Attending at department/faculty/university meetings: The percentage of times that the instructor as memberships attended at the department/ faculty/ university meeting.

Attending at department/faculty/university meetings	Score
Little (0 – 24 %)	0
Some (25 – 49 %)	1
Many (50 – 74 %)	2
Always (75 – 99 %)	3
All (100%)	4

6.4) Keeps systematic documentary evidence of work undertaken: The level of keeping systematic documentary evidence of his work.

Keeps systematic documentary evidence of work undertaken	Score
Never keep	0
Keep some but no system	1
Keep many but not all in system	2
Keep many and all in system	3
Keep all and all in system	4

6.5) Number of times attended academic training/seminars/conferences: Number of times that instructor attended academic training, meetings, or conferences inside or outside the university.

Number of times attended academic training/seminars/conferences	Score
None	0
1 - 2	1
3 - 4	2
5 - 6	3
More than 6	4

6.6) Implementation of the knowledge gained from training/seminars/ conferences: The level of Implementation of the knowledge that the instructor gained from trainings, seminars, or conferences.

Implementation of the knowledge	Score
Never	0
Report to faculty	1
Transfer knowledge to peers	2
Utilize some knowledge	3
Utilize knowledge that useful for teaching and research	4

Appendix D

Test of Linearity or Non-linearity

Table D.1 Test of Linearity and Non-linearity for TU instructors' performance scores (L=Linear, N=Non-linear)

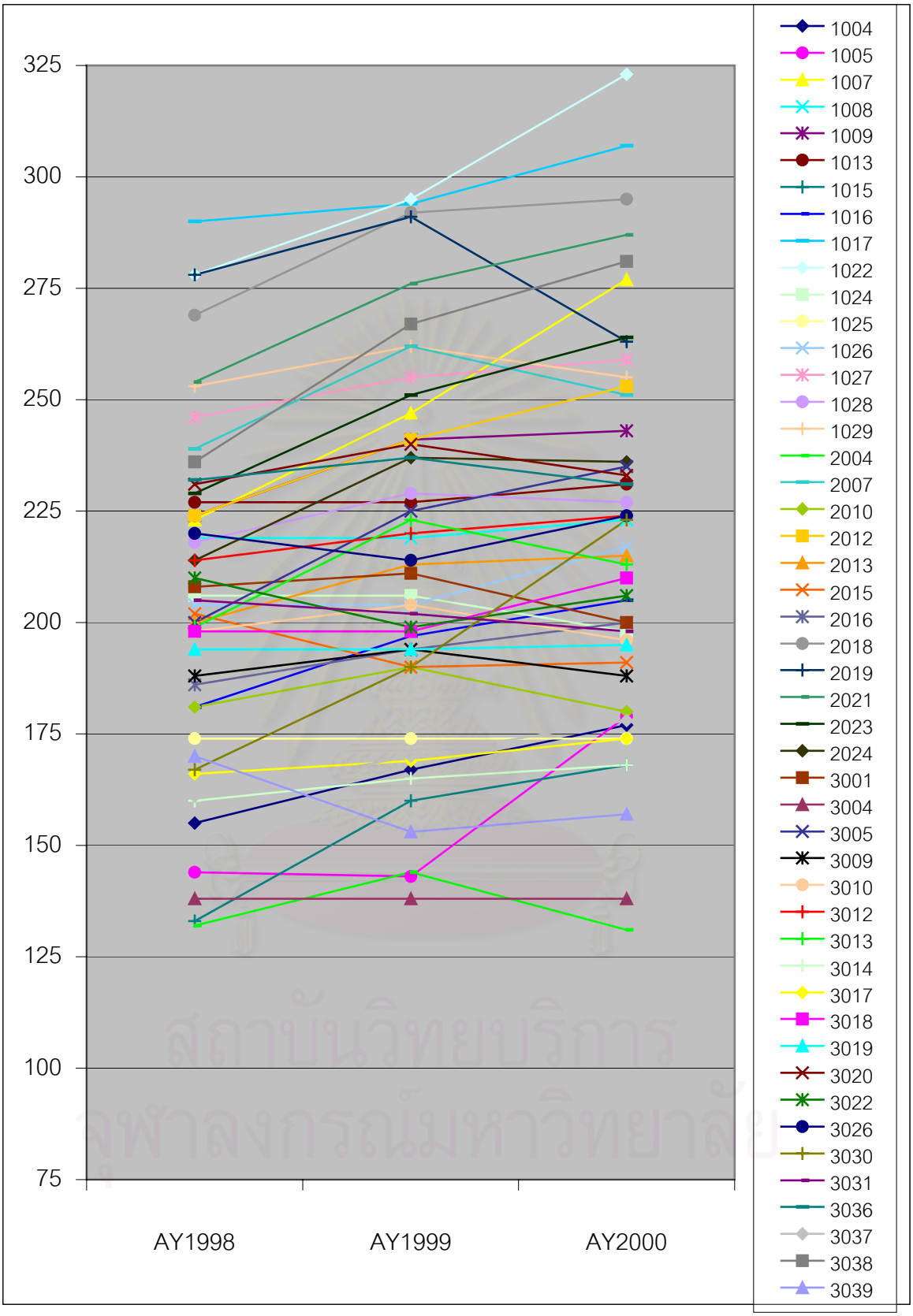
Code	Performance			Slope		Type
	1998	1999	2000	M_1	M_2	
1004	155	167	177	12.0	11.0	Q
1005	144	143	179	-1.0	17.5	Q
1007	223	247	277	24.0	27.0	Q
1008	219	219	223	0.0	2.0	Q
1009	224	241	243	17.0	9.5	Q
1013	227	227	231	0.0	2.0	Q
1015	133	160	168	27.0	17.5	Q
1016	181	197	205	16.0	12.0	Q
1017	290	294	307	4.0	8.5	Q
1022	278	295	323	17.0	22.5	Q
1024	206	206	198	0.0	-4.0	Q
1025	174	174	174	0.0	0.0	L
1026	204	204	217	0.0	6.5	Q
1027	246	255	259	9.0	6.5	Q
1028	218	229	227	11.0	4.5	Q
1029	253	262	255	9.0	1.0	Q
2004	132	144	131	12.0	-0.5	Q
2007	239	262	251	23.0	6.0	Q
2010	181	190	180	9.0	-0.5	Q
2012	224	241	253	17.0	14.5	Q
2013	200	213	215	13.0	7.5	Q
2015	202	190	191	-12.0	-5.5	Q
2016	186	194	200	8.0	7.0	Q
2018	269	292	295	23.0	13.0	Q
2019	278	291	263	13.0	-7.5	Q
2021	254	276	287	22.0	16.5	Q
2023	229	251	264	22.0	17.5	Q
2024	214	237	236	23.0	11.0	Q
3001	208	211	200	3.0	-4.0	Q
3004	138	138	138	0.0	0.0	L
3005	200	225	235	25.0	17.5	Q
3009	188	194	188	6.0	0.0	Q
3010	198	204	196	6.0	-1.0	Q
3012	214	220	224	6.0	5.0	Q
3013	199	223	213	24.0	7.0	Q
3014	160	165	168	5.0	4.0	Q
3017	166	169	174	3.0	4.0	Q
3018	198	198	210	0.0	6.0	Q
3019	194	194	195	0.0	0.5	Q
3020	231	240	233	9.0	1.0	Q
3022	210	199	206	-11.0	-2.0	Q
3026	220	214	224	-6.0	2.0	Q
3030	167	190	223	23.0	28.0	Q
3031	205	202	198	-3.0	-3.5	Q
3036	232	237	231	5.0	-0.5	Q
3037	161	170	178	9.0	8.5	Q
3038	236	267	281	31.0	22.5	Q
3039	170	153	157	-17.0	-6.5	Q
4001	260	269	267	9.0	3.5	Q
4003	212	210	231	-2.0	9.5	Q
4004	249	242	256	-7.0	3.5	Q
4005	225	248	255	23.0	15.0	Q
4006	205	206	212	1.0	3.5	Q
4007	205	226	219	21.0	7.0	Q
4009	222	224	237	2.0	7.5	Q
4011	210	216	217	6.0	3.5	Q
4012	207	208	210	1.0	1.5	Q
4013	257	271	278	14.0	10.5	Q
4014	238	227	219	-11.0	-9.5	Q
4015	204	222	253	18.0	24.5	Q
4017	226	235	254	9.0	14.0	Q
4018	95	225	222	130	63.5	Q
4019	164	213	259	49.0	47.5	Q
4020	259	268	281	9.0	11.0	Q
4021	172	181	190	9.0	9.0	L
4022	197	216	233	19.0	18.0	Q
4023	272	276	280	4.0	4.0	L
5002	196	192	193	-4.0	-1.5	Q
5003	175	187	190	12.0	7.5	Q
5006	198	204	245	6.0	23.5	Q
5008	209	224	235	15.0	13.0	Q
5010	226	228	239	2.0	6.5	Q
5011	233	250	252	17.0	9.5	Q
5012	237	244	251	7.0	7.0	L
5013	227	225	230	-2.0	1.5	Q
5015	165	161	165	-4.0	0.0	Q
5016	209	213	232	4.0	11.5	Q
5017	199	206	206	7.0	3.5	Q
5020	207	202	221	-5.0	7.0	Q
5021	223	223	230	0.0	3.5	Q
5023	188	193	203	5.0	7.5	Q
5034	216	219	234	3.0	9.0	Q
5035	185	184	184	-1.0	-0.5	Q
5036	220	221	219	1.0	-0.5	Q
5037	264	265	262	1.0	-1.0	Q
5039	242	250	240	8.0	-1.0	Q
5040	271	260	257	-11.0	-7.0	Q
5041	241	250	270	9.0	14.5	Q
5042	172	173	183	1.0	5.5	Q
5043	215	217	219	2.0	2.0	L
5044	166	197	209	31.0	21.5	Q
5045	226	240	239	14.0	6.5	Q
5046	224	229	242	5.0	9.0	Q
5047	177	178	193	1.0	8.0	Q
5048	178	181	203	3.0	12.5	Q
5050	217	240	251	23.0	17.0	Q
5052	242	234	242	-8.0	0.0	Q
5053	219	232	237	13.0	9.0	Q
5054	253	266	282	13.0	14.5	Q
5055	213	223	254	10.0	20.5	Q
5056	238	244	247	6.0	4.5	Q
5057	270	283	283	13.0	6.5	Q
5058	172	201	202	29.0	15.0	Q
5059	200	205	210	5.0	5.0	L
5060	170	171	167	1.0	-1.5	Q
5061	212	223	226	11.0	7.0	Q
5063	181	178	190	-3.0	4.5	Q
5064	176	201	211	25.0	17.5	Q
5065	134	186	206	52.0	36.0	Q
5066	174	181	186	7.0	6.0	Q
5067	134	188	179	54.0	22.5	Q
5069	256	270	283	14.0	13.5	Q
6002	222	241	243	19.0	10.5	Q
6003	267	269	267	2.0	0.0	Q
6010	203	217	228	14.0	12.5	Q
6011	175	185	174	10.0	-0.5	Q
6012	286	273	272	-13.0	-7.0	Q
6013	230	233	248	3.0	9.0	Q
6014	168	162	158	-6.0	-5.0	Q
6017	217	223	224	6.0	3.5	Q
6020	166	188	180	22.0	7.0	Q
6021	181	176	174	-5.0	-3.5	Q
6022	301	285	286	-16.0	-7.5	Q
6025	238	265	250	27.0	6.0	Q
7001	201	201	214	0.0	6.5	Q
7003	192	191	195	-1.0	1.5	Q
7004	194	208	225	14.0	15.5	Q
7006	168	172	172	4.0	2.0	Q
7012	241	244	254	3.0	6.5	Q
7014	158	175	198	17.0	20.0	Q
7019	186	213	217	27.0	15.5	Q
7021	188	175	182	-13.0	-3.0	Q

Code	Performance			Slope		Type
	1998	1999	2000	M_1	M_2	
9004	159	182	193	23.0	17.0	Q
9005	142	123	141	-19.0	-0.5	Q
9010	152	152	155	0.0	1.5	Q
9011	240	241	249	1.0	4.5	Q
9013	168	174	174	6.0	3.0	Q
9014	175	198	201	23.0	13.0	Q
9016	161	203	230	42.0	34.5	Q
9019	169	173	173	4.0	2.0	Q
9023	199	210	214	11.0	7.5	Q
9024	145	174	188	29.0	21.5	Q
9025	227	231	233	4.0	3.0	Q
9026	156	209	235	53.0	39.5	Q
9027	218	255	269	37.0	25.5	Q
9028	236	241	253	5.0	8.5	Q
9029	181	209	236	28.0	27.5	Q
9030	206	220	240	14.0	17.0	Q
9031	150	148	182	-2.0	16.0	Q
9032	193	218	223	25.0	15.0	Q
9035	218	239	244	21.0	13.0	Q
9036	183	204	212	21.0	14.5	Q
9037	142	153	153	11.0	5.5	Q
9038	93	102	106	9.0	6.5	Q
9039	143	186	186	43.0	21.5	Q
9041	156	177	185	21.0	14.5	Q
9042	155	174	194	19.0	19.5	Q
9043	158	178	208	20.0	25.0	Q
9045	191	192	203	1.0	6.0	Q
9046	149	134	152	-15.0	1.5	Q
9047	216	230	237	14.0	10.5	Q
9048	149	162	163	13.0	7.0	Q
9055	187	206	212	19.0	12.5	Q
9056	148	210	235	62.0	43.5	Q
9058	230	236	238	6.0	4.0	Q
9059	156	221	247	65.0	45.5	Q
9064	234	239	251	5.0	8.5	Q
9065	181	228	222	47.0	20.5	Q
9066	236	263	263	27.0	13.5	Q
9067	202	219	231	17.0	14.5	Q
9068	152	148	152	-4.0	0.0	Q
10002	168	223	216	55.0	24.0	Q
10005	214	233	236	19.0	11.0	Q
10006	108	130	146	22.0	19.0	Q
10007	160	209	250	49.0	45.0	Q
10008	185	202	209	17.0	12.0	Q
10009	181	195	197	14.0	8.0	Q
10010	127	128	137	1.0	5.0	Q
10011	175	194	192	19.0	8.5	Q
10014	91	95	95	4.0	2.0	Q
10015	232	240	237	8.0	2.5	Q
10016	210	237	245	27.0	17.5	Q
10017	199	205	213	6.0	7.0	Q
10019	183	187	191	4.0	4.0	L
10020	167	168	168	1.0	0.5	Q
10022	114	131	146	17.0	16.0	Q
11001	154	190	226	36.0	36.0	L
11002	250	242	241	-8.0	-4.5	Q
11005	96	117	148	21.0	26.0	Q
11006	172	183	185	11.0	6.5	Q
11007	182	192	209	10.0	13.5	Q
11008	284	289	290	5.0	3.0	Q
11011	195	188	191	-7.0	-2.0	Q
11012	121	165	205	44.0	42.0	Q
11013	167	161	191	-6.0	12.0	Q
11017	212	219	228	7.0	8.0	Q
11018	160	205	229	45.0	34.5	Q
11020	172	208	243	36.0	35.5	Q
11026	191	223	238	32.0	23.5	Q
11028	230	242	242	12.0	6.0	Q
11033	180	213	236	33.0	28.0	Q

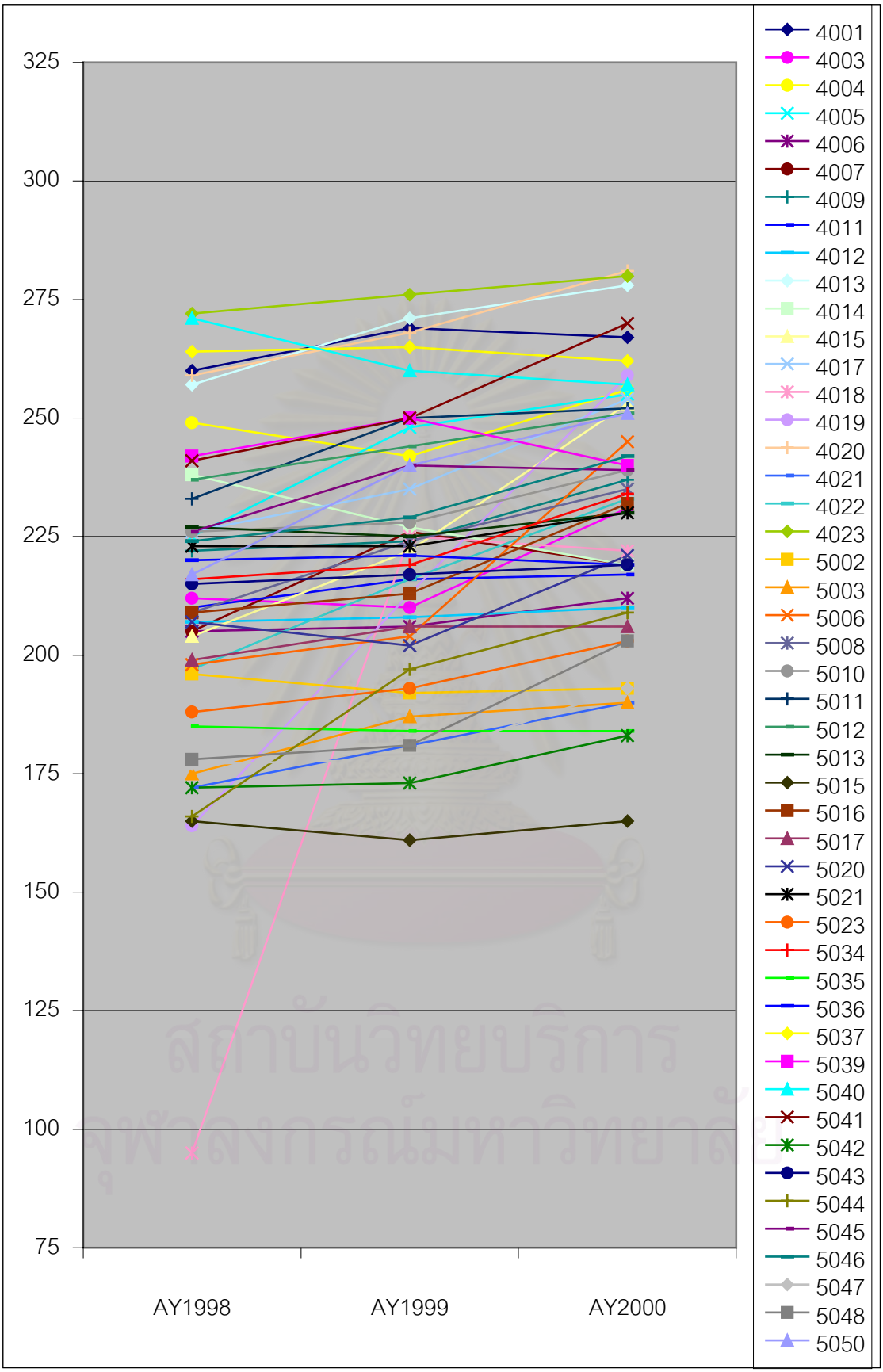
Code	Performance			Slope		Type
	1998	1999	2000	M_1	M_2	
11035	131	137	139	6.0	4.0	Q
11036	220	212	214	-8.0	-3.0	Q
11037	246	227	222	-19.0	-12.0	Q
11040	150	177	210	27.0	30.0	Q
11041	177	204	222	27.0	22.5	Q
11046	254	254	271	0.0	8.5	Q
11047	201	164	163	-37.0	-19.0	Q
11048	242	268	286	26.0	22.0	Q
11059	221	240	261	19.0	20.0	Q
11060	207	244	244	37.0	18.5	Q
11063	250	250	254	0.0	2.0	Q
11064	189	194	207	5.0	9.0	Q
11065	238	248	242	10.0	2.0	Q
11067	75	195	223	120	74.0	Q
12001	217	194	201	-23.0	-8.0	Q
12002	200	197	194	-3.0	-3.0	L
12003	171	206	214	35.0	21.5	Q
12004	207	227	219	20.0	6.0	Q
12005	184	205	215	21.0	15.5	Q
12006	182	149	141	-33.0	-20.5	Q
12007	154	191	203	37.0	24.5	Q
13001	141	194	222	53.0	40.5	Q
13002	102	187	207	85.0	52.5	Q
13003	92	157	189	65.0	48.5	Q
13004	139	169	191	30.0	26.0	Q
13005	222	221	205	-1.0	-8.5	Q
13006	96	178	199	82.0	51.5	Q
13007	98	154	192	56.0	47.0	Q
14001	271	276	278	5.0	3.5	Q
14002	273	266	268	-7.0	-2.5	Q
14003	181	203	220	22.0	19.5	Q
14005	211	241	246	30.0	17.5	Q
14007	198	242	250	44.0	26.0	Q
14008	267	272	280	5.0	6.5	Q
14009	225	241	240	16.0	7.5	Q
14010	169	193	206	24.0	18.5	Q
14011	224	241	277	17.0	26.5	Q

Table D.2 Test of Linearity and Non-linearity for SU instructors' performance scores (L=Linear, N=Non-linear)

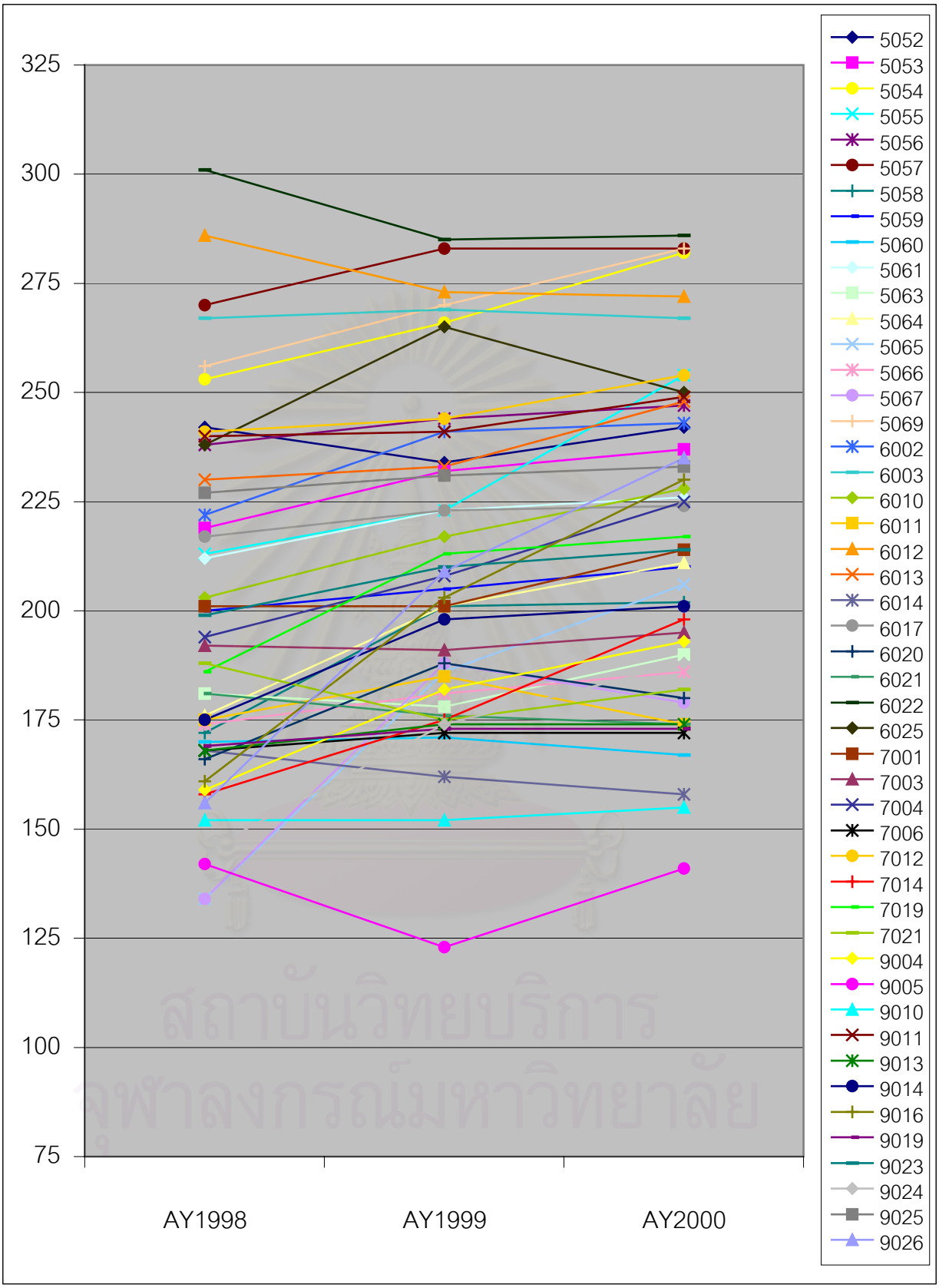
Code	Performance			Slope		Type
	1998	1999	2000	M_1	M_2	
101	189	196	202	7.0	6.5	Q
102	275	270	282	-5.0	3.5	Q
103	175	218	270	43.0	47.5	Q
104	246	256	277	10.0	15.5	Q
106	212	218	252	6.0	20.0	Q
109	198	223	238	25.0	20.0	Q
211	203	259	274	56.0	35.5	Q
215	181	200	200	19.0	9.5	Q
221	203	242	261	39.0	29.0	Q
227	159	220	250	61.0	45.5	Q
228	242	243	244	1.0	1.0	L
229	228	249	269	21.0	20.5	Q
231	207	222	201	15.0	-3.0	Q
232	90	132	153	42.0	31.5	Q
233	204	221	231	17.0	13.5	Q
234	100	158	194	58.0	47.0	Q
235	241	273	294	32.0	26.5	Q
237	188	222	247	34.0	29.5	Q
238	194	248	254	54.0	30.0	Q
239	144	164	169	20.0	12.5	Q
240	163	189	208	26.0	22.5	Q
241	145	152	157	7.0	6.0	Q
242	157	221	251	64.0	47.0	Q
301	134	200	213	66.0	39.5	Q
302	219	263	267	44.0	24.0	Q
305	263	271	279	8.0	8.0	L
306	218	221	225	3.0	3.5	Q
308	97	97	98	0.0	0.5	Q
310	169	208	224	39.0	27.5	Q
311	265	269	266	4.0	0.5	Q
312	202	213	241	11.0	19.5	Q
313	272	241	305	-31.0	16.5	Q
401	235	272	279	37.0	22.0	Q
402	144	169	193	25.0	24.5	Q
403	231	260	276	29.0	22.5	Q
502	265	273	289	8.0	12.0	Q
503	205	205	216	0.0	5.5	Q
505	282	298	316	16.0	17.0	Q
506	200	218	244	18.0	22.0	Q
601	134	219	247	85.0	56.5	Q
610	127	131	146	4.0	9.5	Q
611	175	198	219	23.0	22.0	Q
612	179	199	215	20.0	18.0	Q
617	191	208	217	17.0	13.0	Q
619	155	169	151	14.0	-2.0	Q
620	129	175	190	46.0	30.5	Q
621	165	178	188	13.0	11.5	Q
622	112	173	219	61.0	53.5	Q
623	193	240	262	47.0	34.5	Q
624	154	196	226	42.0	36.0	Q
625	169	215	221	46.0	26.0	Q
626	173	198	212	25.0	19.5	Q



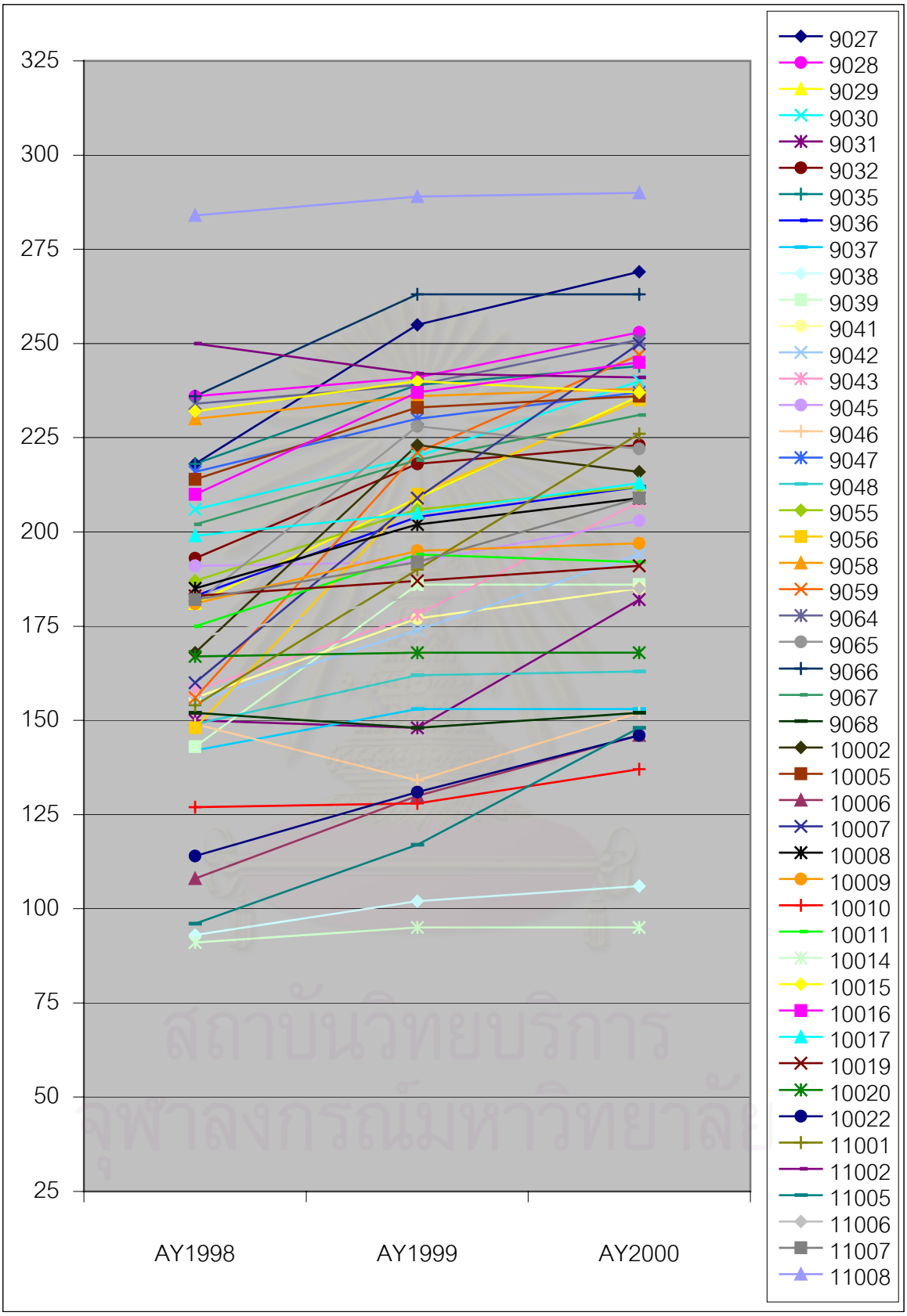
Graph D.1 TU Instructors' Performance Scores in Academic Years 1998, 1999 and 2000



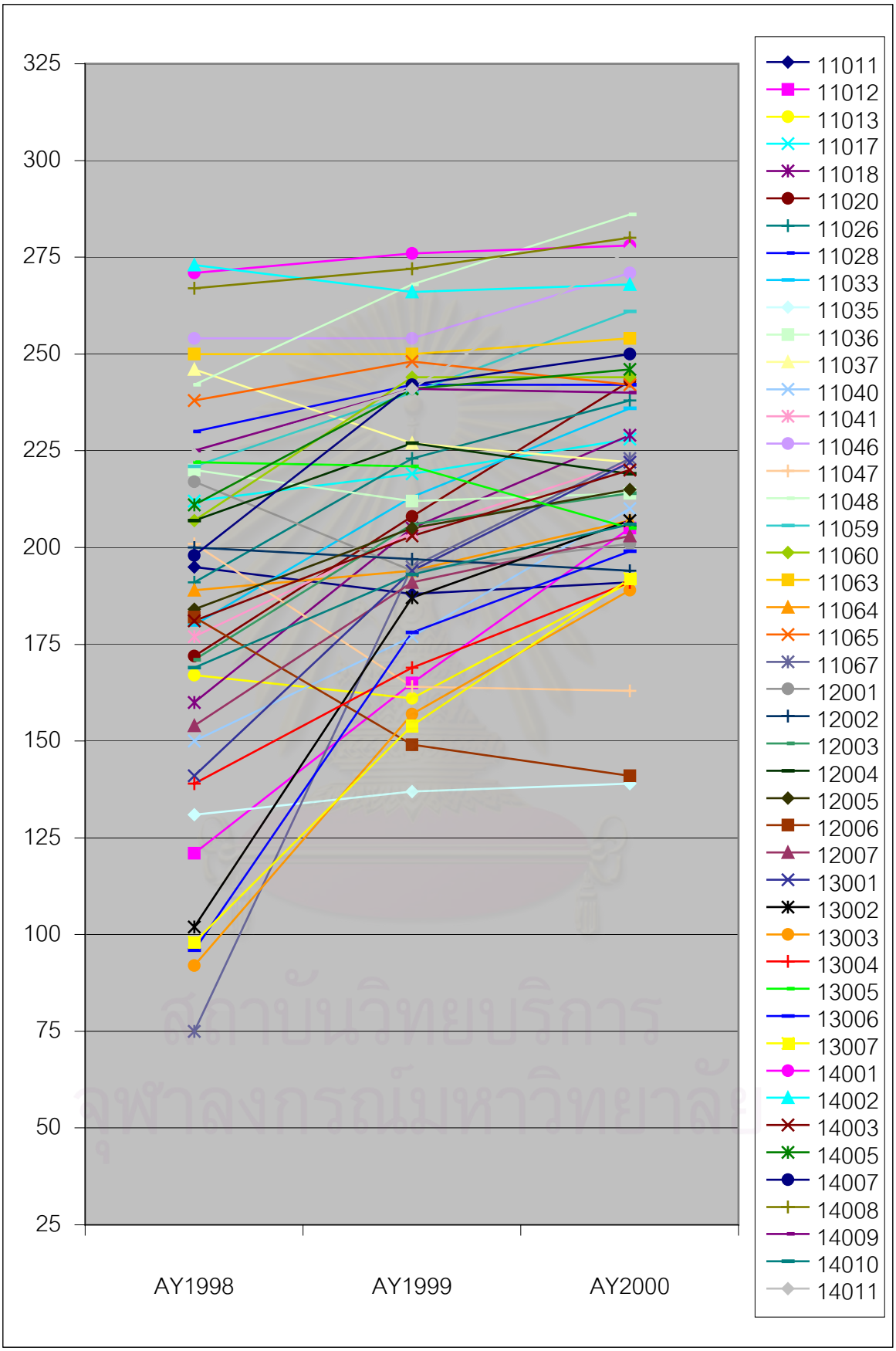
Graph D.1 TU Instructors' Performance Scores in Academic Years 1998, 1999 and 2000 (Continue)



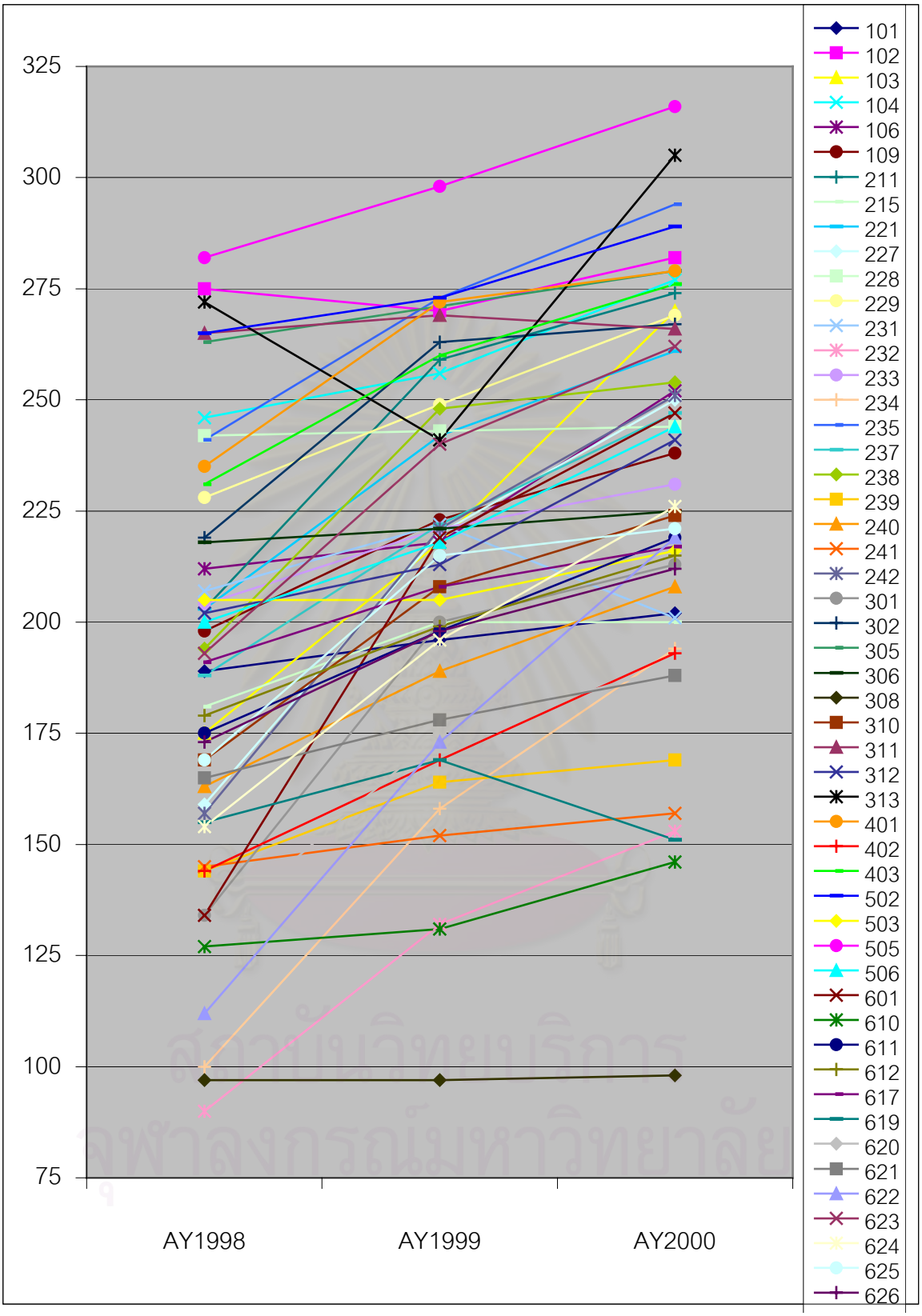
Graph D.1 TU Instructors' Performance Scores in Academic Years 1998, 1999 and 2000 (Continue)



Graph D.1 TU Instructors' Performance Scores in Academic Years 1998, 1999 and 2000 (Continue)



Graph D.1 TU Instructors' Performance Scores in Academic Years 1998, 1999 and 2000 (Continue)



Graph D.2 SU Instructors' Performance Scores in Academic Years 1998, 1999 and 2000

Appendix E

Performance Scores & Performance Change Scores

Table E.1 Performance Scores and Performance Change Scores of TU instructors (F_1=change scores calculated by researcher's formula, F_2=change scores calculated by Bryk & Raudenbush's formula)

Code	Performance			Change	
	1998	1999	2000	F_1 (Max=150)	F_2 (Max=200)
1004	155	167	177	9.91	11.41
1005	144	143	179	15.01	16.78
1007	223	247	277	29.45	23.61
1008	219	219	223	2.10	3.6
1009	224	241	243	10.11	9.54
1013	227	227	231	2.14	3.52
1015	133	160	168	14.84	16.8
1016	181	197	205	11.59	11.95
1017	290	294	307	10.48	8.15
1022	278	295	323	27.48	19.49
1024	206	206	198	-3.94	-1.11
1025	174	174	174	0.00	2.41
1026	204	204	217	6.57	7.36
1027	246	255	259	7.29	6.94
1028	218	229	227	4.68	5.59
1029	253	262	255	1.05	2.45
2004	132	144	131	-0.39	2.37
2007	239	262	251	6.27	6.56
2010	181	190	180	-0.41	1.92
2012	224	241	253	15.63	13.56
2013	200	213	215	7.53	8.16
2015	202	190	191	-5.46	-2.24
2016	186	194	200	6.82	7.91
2018	269	292	295	15.33	11.91
2019	278	291	263	-5.58	-4.64
2021	254	276	287	19.03	14.87
2023	229	251	264	19.13	15.91
2024	214	237	236	11.36	10.82
3001	208	211	200	-3.88	-1.14
3004	138	138	138	0.00	2.75
3005	200	225	235	17.81	16.17
3009	188	194	188	0.00	2.26
3010	198	204	196	-0.95	1.36
3012	214	220	224	5.21	6.04
3013	199	223	213	6.68	7.74
3014	160	165	168	3.62	5.75
3017	166	169	174	3.68	5.7
3018	198	198	210	5.97	7.02
3019	194	194	195	0.49	2.63
3020	231	240	233	1.00	2.65
3022	210	199	206	-1.85	0.5
3026	220	214	224	2.03	3.61
3030	167	190	223	26.64	24.94
3031	205	202	198	-3.47	-0.69
3036	232	237	231	-0.42	1.45
3037	161	170	178	7.76	9.35
3038	236	267	281	25.07	19.84
3039	170	153	157	-6.90	-2.73
4001	260	269	267	4.00	4.39
4003	212	210	231	9.73	9.71
4004	249	242	256	3.81	4.55
4005	225	248	255	16.14	13.93
4006	205	206	212	3.55	4.94
4007	205	226	219	6.90	7.69
4009	222	224	237	7.95	7.99
4011	210	216	217	3.60	4.87
4012	207	208	210	1.53	3.31
4013	257	271	278	12.12	10.04
4014	238	227	219	-8.71	-5.8
4015	204	222	253	25.44	21.79
4017	226	235	254	15.12	13.16
4018	95	225	222	46.34	53.85
4019	164	213	259	46.04	40.58
4020	259	268	281	12.77	10.44
4021	172	181	190	8.47	9.65
4022	197	216	233	18.27	16.62
4023	272	276	280	4.74	4.7
5002	196	192	193	-1.51	1.01
5003	175	187	190	7.08	8.4
5006	198	204	245	23.69	21.08
5008	209	224	235	13.50	12.5
5010	226	228	239	6.95	7.15
5011	233	250	252	10.33	9.45
5012	237	244	251	7.71	7.43
5013	227	225	230	1.59	3.13
5015	165	161	165	0.00	2.51
5016	209	213	232	11.85	11.33
5017	199	206	206	3.49	4.97
5020	207	202	221	7.00	7.75
5021	223	223	230	3.70	4.77
5023	188	193	203	7.34	8.3
5034	216	219	234	9.42	9.25
5035	185	184	184	-0.52	1.91
5036	220	221	219	-0.47	1.57
5037	264	265	262	-0.84	0.76
5039	242	250	240	-0.80	0.94
5040	271	260	257	-5.80	-4.1
5041	241	250	270	16.21	13.42
5042	172	173	183	5.13	6.86
5043	215	217	219	2.08	3.63
5044	166	197	209	20.14	19.69
5045	226	240	239	6.91	7.11
5046	224	229	242	9.62	9.17
5047	177	178	193	7.56	8.82
5048	178	181	203	11.90	12.43
5050	217	240	251	18.04	15.62
5052	242	234	242	0.00	1.8
5053	219	232	237	9.51	9.19
5054	253	266	282	16.68	13.3
5055	213	223	254	21.55	18.51
5056	238	244	247	4.95	5.41
5057	270	283	283	7.64	6.7
5058	172	201	202	13.99	14.41
5059	200	205	210	5.03	6.18
5060	170	171	167	-1.61	1.24
5061	212	223	226	7.25	7.65
5063	181	178	190	4.24	5.98
5064	176	201	211	16.76	16.4
5065	134	186	206	31.36	31.6
5066	174	181	186	5.65	7.22
5067	134	188	179	18.18	20.73
5069	256	270	283	15.62	12.46
6002	222	241	243	11.13	10.35
6003	267	269	267	0.00	1.53
6010	203	217	228	12.79	12.16
6011	175	185	174	-0.39	1.97
6012	286	273	272	-5.51	-4.23
6013	230	233	248	9.73	9.12
6014	168	162	158	-5.43	-1.54
6017	217	223	224	3.66	4.81
6020	166	188	180	6.19	8.05
6021	181	176	174	-3.68	-0.45
6022	301	285	286	-5.59	-4.77
6025	238	265	250	6.06	6.56
7001	201	201	214	6.52	7.39
7003	192	191	195	1.47	3.45
7004	194	208	225	15.57	14.66
7006	168	172	172	1.84	4.07
7012	241	244	254	7.20	7
7014	158	175	198	18.39	18.61
7019	186	213	217	15.09	14.69
7021	188	175	182	-2.98	-0.09

Code	Performance			Change	
	1998	1999	2000	F_1 (Max=150)	F_2 (Max=200)
9004	159	182	193	15.57	16.16
9005	142	123	141	-0.15	2.37
9010	152	152	155	1.32	3.83
9011	240	241	249	4.96	5.41
9013	168	174	174	2.76	4.87
9014	175	198	201	12.27	12.79
9016	161	203	230	32.55	30.17
9019	169	173	173	1.85	4.06
9023	199	210	214	7.54	8.18
9024	145	174	188	19.05	19.9
9025	227	231	233	3.21	4.32
9026	156	209	235	36.88	34.21
9027	218	255	269	27.30	22.41
9028	236	241	253	9.34	8.65
9029	181	209	236	27.14	24.39
9030	206	220	240	17.61	15.75
9031	150	148	182	13.92	15.52
9032	193	218	223	14.89	14.23
9035	218	239	244	13.72	12.4
9036	183	204	212	14.09	13.93
9037	142	153	153	4.70	7.11
9038	93	102	106	4.81	8.38
9039	143	186	186	18.44	19.87
9041	156	177	185	13.12	14.19
9042	155	174	194	17.78	18.23
9043	158	178	208	23.13	22.62
9045	191	192	203	5.88	7.08
9046	149	134	152	0.97	3.9
9047	216	230	237	11.04	10.42
9048	149	162	163	6.12	8.24
9055	187	206	212	12.24	12.29
9056	148	210	235	39.78	37.47
9058	230	236	238	4.32	5.09
9059	156	221	247	42.61	39
9064	234	239	251	9.30	8.67
9065	181	228	222	19.17	18.7
9066	236	263	263	14.72	12.61
9067	202	219	231	14.83	13.77
9068	152	148	152	0.00	2.63
10002	168	223	216	21.60	21.61
10005	214	233	236	11.46	10.83
10006	108	130	146	15.07	18.26
10007	160	209	250	43.01	38.61
10008	185	202	209	11.70	11.91
10009	181	195	197	7.66	8.74
10010	127	128	137	4.10	6.88
10011	175	194	192	7.92	9.19
10014	91	95	95	1.46	4.79
10015	232	240	237	2.67	3.85
10016	210	237	245	18.21	16.07
10017	199	205	213	7.04	7.79
10019	183	187	191	3.85	5.54
10020	167	168	168	0.46	2.88
10022	114	131	146	12.88	15.8
11001	154	190	226	33.48	31.46
11002	250	242	241	-3.95	-1.9
11005	96	117	148	20.05	24.01
11006	172	183	185	6.07	7.63
11007	182	192	209	13.11	13.17
11008	284	289	290	3.64	3.78
11011	195	188	191	-2.00	0.63
11012	121	165	205	35.91	36.58
11013	167	161	191	10.79	12.16
11017	212	219	228	8.32	8.47
11018	160	205	229	32.40	30.17
11020	172	208	243	34.59	30.89
11026	191	223	238	23.57	21.07
11028	230	242	242	6.45	6.68
11033	180	213	236	27.55	24.79

Code	Performance			Change	
	1998	1999	2000	F_1 (Max=150)	F_2 (Max=200)
11035	131	137	139	3.33	6.02
11036	220	212	214	-2.83	-0.41
11037	246	227	222	-10.74	-7.86
11040	150	177	210	27.36	26.7
11041	177	204	222	21.81	20.41
11046	254	254	271	9.65	8.5
11047	201	164	163	-19.00	-13.02
11048	242	268	286	24.90	19.4
11059	221	240	261	21.55	18.01
11060	207	244	244	18.82	16.88
11063	250	250	254	2.25	3.31
11064	189	194	207	8.83	9.5
11065	238	248	242	2.12	3.39
11067	75	195	223	55.08	62.52
12001	217	194	201	-7.46	-4.36
12002	200	197	194	-3.01	-0.24
12003	171	206	214	20.29	19.64
12004	207	227	219	5.91	6.87
12005	184	205	215	15.14	14.73
12006	182	149	141	-21.75	-14.06
12007	154	191	203	22.24	22.2
13001	141	194	222	36.38	35.15
13002	102	187	207	41.76	45.08
13003	92	157	189	38.01	42.01
13004	139	169	191	22.86	23.57
13005	222	221	205	-8.05	-4.88
13006	96	178	199	40.26	44.34
13007	98	154	192	37.68	40.78
14001	271	276	278	4.13	4.3
14002	273	266	268	-2.03	-0.51
14003	181	203	220	19.04	17.97
14005	211	241	246	18.17	16.06
14007	198	242	250	26.31	22.97
14008	267	272	280	7.64	6.75
14009	225	241	240	7.95	7.92
14010	169	193	206	17.46	17.27
14011	224	241	277	28.86	23.22

Table E.1 Performance Scores and Performance Change Scores of SU instructors (F_1=change scores calculated by researcher's formula, F_2=change scores calculated by Bryk & Raudenbush's formula)

Code	Performance			Change	
	1998	1999	2000	F_1 (Max=150)	F_2 (Max=200)
101	189	196	202	6.37	9.95
102	275	270	282	4.08	7.12
103	175	218	270	47.33	41.54
104	246	256	277	17.55	16.51
106	212	218	252	20.86	20.21
109	198	223	238	20.37	20.25
211	203	259	274	36.82	32.09
215	181	200	200	9.05	12.29
221	203	242	261	30.14	27.12
227	159	220	250	43.12	40.06
228	242	243	244	1.11	5.39
229	228	249	269	22.46	20.46
231	207	222	201	-2.55	2.51
232	90	132	153	23.94	29.76
233	204	221	231	13.85	15.23
234	100	158	194	37.86	41.6
235	241	273	294	30.06	24.97
237	188	222	247	29.68	27.61
238	194	248	254	29.96	27.92
239	144	164	169	10.88	14.83
240	163	189	208	21.04	22.4
241	145	152	157	5.22	9.85
242	157	221	251	44.35	41.22
301	134	200	213	34.06	35.59
302	219	263	267	25.36	23.16
305	263	271	279	9.34	10.64
306	218	221	225	3.67	7.47
308	97	97	98	0.37	5.94
310	169	208	224	26.15	26.18
311	265	269	266	0.57	4.86
312	202	213	241	19.98	19.87
313	272	241	305	16.99	17.21
401	235	272	279	24.25	21.53
402	144	169	193	21.82	24.06
403	231	260	276	24.82	21.96
502	265	273	289	14.11	13.7
503	205	205	216	5.57	9.1
505	282	298	316	20.85	17.43
506	200	218	244	22.59	21.8
601	134	219	247	50.15	48.64
610	127	131	146	7.84	12.67
611	175	198	219	21.23	21.95
612	179	199	215	17.46	18.85
617	191	208	217	12.90	14.93
619	155	169	151	-1.91	3.61
620	129	175	190	25.95	28.73
621	165	178	188	10.66	13.95
622	112	173	219	45.24	46.52
623	193	240	262	35.19	31.4
624	154	196	226	33.44	32.82
625	169	215	221	24.33	25.01
626	173	198	212	18.62	20.03

Biography

I am Jitlekha Teerajarmorn, the Ph.D. Thailand Research Fund scholarship student in Educational Measurement and Evaluation program since the first semester of the academic year 1998. I graduated with Bachelor's degrees (first class honor) in Secondary Education, Thai and Guidance and Counseling.

I was born in Phuket on February 12, 1977. I was invited to undertake a student attachment with the Research and Evaluation Division of the University of Cambridge Local Examinations Syndicate, England and also invited to cooperate with Professor in Newcastle University Upon Tyne. Moreover, I was teacher assistance in Statistical Course and research assistance of Professor Dr.Utumporn Jamornmann.