Using Faculty Evaluation to Improve Teaching Quality:

A Longitudinal Case Study of Higher Education in Southeast Asia

Philip Hallinger[1]

Hong Kong Institute of Education

Hallinger@gmail.com

Keywords: faculty evaluation, performance assessment, performance and reward, higher education, organizational change

[1] The author wishes to acknowledge the assistance of Assistant Professor Lu Jiafang and Research Assistant Li Fang in conducting the data analysis for this study.

Abstract

This paper presents a case study of implementation of a new system of faculty teaching evaluation at a graduate school of business in Thailand. This paper employs a case study methodology to describe the implementation of a new system of teacher performance evaluation over a seven-year period. The paper describes the rationale behind the development of the College's faculty evaluation system and its role in the College's quality improvement program. Quantitative data are presented that show trends in the improvement of teaching in the College over the even-year period.

Higher education institutions find themselves in an environment of increased accountability for the quality of teaching and learning they provide to students (Astin, 1999; Jones, 2009). Although the use of accountability tools in higher education began as a largely 'Western' phenomenon, in recent years their adoption has become global in scale. The impact of quality assurance processes driven by government and accreditation agencies has been further accentuated in recent years by the dissemination of World University Rankings. Today, universities throughout the world find themselves in a race both to prove their merit and 'keep up' with local and global competitors. This has, in turn, fostered the use of new management tools aimed at strengthening performance on key performance indicators associated with university quality.

This report tells the story of how one institution, a graduate school of business (GSB), located in Thailand, responded to the quality challenge. This case study examines the design and implementation of a system of faculty evaluation built upon strategic human resource management principles of performance assessment and reward (Lawler, 2008). The paper presents the results of a longitudinal, non-experimental evaluation of the GSB's strategy for improving the quality of teaching. More specifically, the purposes of this paper are to:

1. Discuss the design of a system of faculty evaluation within a broader program of quality improvement;
2. Present results on the impact of the new system of faculty evaluation on instructor effectiveness and faculty turnover;
3. Explore the complexities and test the potential of employing performance assessment and reward in higher education, with an emphasis on universities in developing countries such as in Southeast Asia.

The value of this study lies first in assessing the potential that systems of performance assessment and reward can offer in higher education. While the contextual conditions of any case study limit the generalizability of findings, we suggest at the outset that there would be few such 'pure' implementations of faculty evaluation in universities internationally. Thus, even the uniqueness of this case study may offer insight into the conditions necessary to achieve positive effects of performance assessment and reward in higher education.

Second, Asian universities are struggling to meet the challenges of quality improvement in a global game in which they start with fewer 'chips' on the table. This case study offers data that describe the complexities and possibilities of bringing about sustainable quality improvement in teaching and learning in Southeast Asia. We note that while this region has demonstrated the highest rate of growth in tertiary education over the past decade, it has also experienced an uncertain transition to the use of new methods of teaching (Altbach & Umakoshi, 2004; Hallinger & Bridges, 2007; Walker, Bridges, & Chan, 1996).

Finally, although this case study tells a story grounded in the experience of a single institution, the research employs quantitative analysis of a large longitudinal dataset comprised of student course evaluations. Thus, the story is framed around a substantial body of data collected over a period of seven years. Thus, we suggest that the findings address some, though not all, limitations of case studies.

Prologue

The GSB was only two years old when this story begins in 2000. The college was founded in 1998 as a semi-independent unit in a large, research-oriented, government university in Thailand. In terms of governance, the GSB was responsible on curriculum matters to the University Council of the parent university, but reported to its own separate Board of Trustees on matters of general policy. At the time of this study, the GSB offered the

Master of Management degree (M.M.) taught in English, to 750 students annually in its 'international' program.[i] From its inception, the university's senior management envisioned the GSB as a center of innovation for the delivery of management education in Thailand as well as a resource for improving the management of the university at-large. The GSB's mission to innovate in teaching, learning and management was explicitly facilitated by a hybrid governance structure that freed it from rules imposed by the government system on the parent university.[ii]

The GSB's vision was to offer a personalized, learner-centered education that incorporated global and local perspectives on managing organizations. GSB's stated vision was to 'develop students who are able to apply knowledge effectively in their work and in their lives.' The educational practices implied by this vision were reflected in the curriculum structure and facilities of the college.

GSB facilities were purpose-built to foster student-to-student interaction. All classrooms were equipped with movable tables and chairs, state-of-the-art multi-media projectors, teacher workstations connected to the internet, and stereo sound systems. Unlike competing MBA programs that relied on large class lectures complemented by case teaching in lecture theatres, maximum class size in the GSB was set at 30 students. Even the decision to offer a variety of program specializations under the M.M. degree reflected a desire to give students more choice and flexibility than is typically possible in MBA programs. The combination of mission, vision, curriculum structure, purpose-built classrooms, and small class size was intended to create a new standard in graduate management education in Thailand and differentiate GSB from other local business schools (GSB, 2000).

Despite this seemingly receptive context for innovation in teaching, learning, and management, a quality audit conducted in its third year of operation (i.e., 2000) indicated that the reality was not approaching this ambitious vision. Most instructors kept the tables and

chairs in a traditional classroom seating arrangement. The majority of class time was devoted to teacher-directed instruction broken up by occasional case discussions. Multi-media equipment was used for only the most basic function, electronic delivery of power point slides. Not a single instructor used the technology for multi-media cases, video-enriched content, or access to internet resources during lessons (GSB, 2000). Although the modal practice in Thailand's business education market involved instructors teaching part-time in competitor's programs, the GSB was relying almost 100% on part-time instructors from other universities to deliver its curriculum. Therefore, while a formal curriculum existed on paper, the 'taught curriculum' was a randomly constructed and constantly changing amalgam of courses offered by part-time instructors from other local business schools.

In early 2001, when GSB managers deliberated on the quality audit, they concluded that the college was not organizing to take advantage of its strengths. These included the brand name of the university which was attracting students despite all of the problems, the GSB's location, the best-equipped classroom facilities in Thailand, a healthy budget, freedom to innovate, an entrepreneurial Director, and small class size. Located in a highly competitive market and positioned at upper-mid range in terms of fees, the management team further concluded that GSB's survival would depend on the ability to differentiate its program from other local business schools on the quality of management education**.**

However, threats were imminent. Student dissatisfaction was on the rise. During the quality audit, a group of students had sarcastically referred to the GSB as the "College of Mismanagement" (GSB, 2000). In addition, the quality audit had confirmed the worst fears of the Board of Trustees concerning academic quality. Absent a defensible strategy and immediate execution, the President was poised to change the management team and institute tighter regulation.

It was in this context of immediate threats to survival that the GSB Director formed a new management team to whom he delegated authority for leading the change. The management team developed a multi-pronged strategy to enhance the quality of teaching and learning in the GSB. Although the findings of the quality audit were dire, the management team remained confident both in the GSB's potential strengths as well as in opportunities in the local management education market. In its view, the educational quality of competing management education programs was generally mediocre. The team believed that if a business school could actually organize itself with a reasonable degree of efficiency and demonstrate a capacity to innovate, it could quickly surpass its competitors in attracting high quality students. Of course, if it was easy to accomplish those goals, another institution would already have done so, so the journey towards quality improvement was begun with both an ambitious vision and a sense of trepidation.

With this in mind, the management team formulated a strategy that drew heavily on strategic human resource management (Lawler, 2008). Strategic HR assumes that it is possible to change organizations by shaping behavior through a combination inspiration, evaluation, performance feedback, training/coaching, and reward. We note, however, that this strategy was wholly 'Western' or 'international' in conception. Strategic human resource management was seldom employed by Thai companies, never mind public sector organizations or universities. Indeed the human resource function in most universities still focused almost solely upon personnel processing.

This meant that the management team's strategy would take it through uncharted waters. Among the potential dangers ahead was the response of Thai faculty and staff to a human resource system that was highly instrumental, focusing on goals, performance and results. Nobody could predict the extent to which this approach would be compatible with the norms of Thai culture. For example, in Thailand, it is commonly accepted that 'know who is

more important than know how' when it comes to obtaining a job, getting a raise, or being promoted (Holmes & Tangtongtavy, 1996), much like the system of *guanxi* in China (Park & Luo, 2001). Or as one astute observer of Thai culture noted.

> Responsibility is a proud and cold word, capable of causing abysmal rents in the social fabric and frayed edges of tender feelings. An ethic of compassion, the inculcation of deference to superiors (kreng jai) and an ingrained desire for harmony and familiarity have created a communal security blanket [in Thailand]. . . [Responsibility] signifies 'being the source or cause of something.' It means that one is 'capable of making moral or rational decisions on one's own. . .'' These, in fact, are what it means to be a person in Western terms. Westerners are more committed to taking responsibility because it automatically brings with it the recognition and respect, regardless of consequent profit or loss, accorded to 'real persons'. . . But in Thailand, it is rank and its rewards that give a faint reading of responsibility into the bargain. The honour is not in the responsibility itself (let alone its fulfillment) but in the position that allows it to be borne. (Redmond, 1994, b2)

Thus, the management team selected an instrumental HR strategy knowing that it would run against the grain of cultural norms in Thailand and the parent university. Nonetheless, the strategy was consistent with the GSB's mission to innovate and employ international approaches in the delivery of management education in Thailand. Moreover, these were methods that the GSB instructors taught to students, so it seemed suitable to practice what was preached. The metaphorical goal that guided the team's strategic vision was to 'create a small island in Thailand where competence in higher education would be rewarded'.

We note that initially the quality strategy focused solely upon teaching and learning. Concerns for enhancing research productivity had not yet come into the management team's radar. This was a case of a baby needing to learn to crawl before it could walk. The college relied 100% on student tuition for its survival and had to demonstrate that it could provide a viable management education first. Moreover, in this education context, students were wholly unconcerned with the research productivity of the faculty. They wanted to gain useful knowledge and a graduate qualification from a respected university to enhance their career prospects.

With these points in mind, the strategic thrusts of the GSB's plan for quality improvement were to:

- Communicate a clear vision of learning for the GSB;
- Build instructor capacity in the use of active learning methods;
- Develop a problem-based learning track among its capstone options;
- Reorganize around a more centralized management structure;
- Raise academic standards at the point of admissions and in the classroom;
- Develop a team approach to curriculum design and delivery in selected parts of the program;
- Design and implement a system of faculty performance-based assessment, instructor selection and reward.

The last of these strategic thrusts represents the focus of this report. However, before discussing the design of the faculty evaluation system, we will present the research method employed in this study.

Research Method

This study employed a mixed-methods, longitudinal, non-experimental research design (Creswell, 2007; Yin 2008). Collection of qualitative and quantitative data unfolded

concurrently, term-by-term over a seven-year period. Data analysis for this report sought to provide a long-term assessment of implementation of the faculty evaluation system at the GSB. A key strength of this study's design lies in the longitudinal perspective on change gained through the analysis change in student perceptions of faculty performance and assessment of faculty turnover over a substantial period of time.

Data Collection

This study employed two main categories of data: qualitative data used to construct the narrative on design of the evaluation approach and quantitative data employed to assess the effects of the new faculty evaluation system. Information used to construct the narrative was drawn from two sources. First, the author was the leader of the implementation effort[iii] and maintained personal notes concurrent with implementation. While the author cannot claim objectivity in construction of the narrative, s/he builds the story around quantitative data gathered by the GSB staff and college documents kept during the period of implementation. The author sought to save his/her opinions for the final interpretive section of the paper and then makes these explicit.

As noted, the research draws upon an array of formal and informal documents to construct the historical narrative and make sense of the quantitative results. These included a quality audit conducted in 2000, a report on faculty assessment practices, formative student feedback gathered from Mid-term and end-of-course evaluations, and minutes and reports from the GSB's quality committee.

The main quantitative data collection tool consisted of GSB's Course Evaluation Questionnaire (CEQ) administered to students at the conclusion of each course. Course evaluation questionnaires are subject to a variety of potential problems when used in academic research (Aleamoni, 1999; Scriven, 1988). Nonetheless, a substantial body of research supports the reliability and validity of purposively-designed, systematically

administered course evaluation questionnaires (Aleamoni, 1999; Greenwald & Gilmore, 1997; Harrison, Ryan & Moore, 1996; McKeachie, 1997; Scriven, 1995).

Questionnaire design and procedures for administering and using GSB's CEQ explicitly addressed features that threaten validity (Scriven, 1988, 1995). The 17 item CEQ was designed after a thorough review of scales used by other universities internationally, and in consultation with psychometricians. As recommended in the literature, it contained two global items on course and instructor effectiveness for use in summative decision-making and 15 items that focused on specific teaching related qualifications and behaviors (see Figure 1). The CEQ was administered systematically by GSB academic support staff. Data were entered electronically by an outside company who then provided the academic staff of the GSB with the data file, which was managed by the academic services department.

Although the CEQ contained multiple items and dimensions, in this report we focus on change in instructor effectiveness. This was the most direct indicator of the potential impact of the GSB's quality strategy. Moreover, prior research has found that teacher effects on student ratings are larger than course effects (Marsh, 1981). In this study, rather than relying on a single item, an Instructor Effectiveness scale was defined as a dimension consisting of several items. This dimension was assessed through four items that asked students to rate instructors' knowledge in the subject, preparation for class, clarity of responses to students' questions, and overall rating of the instructor (alpha = .95).

Finally, the study also sought to understand the impact of the performance assessment system on faculty turnover. Turnover rate was assessed by calculating the number of faculty members who left the college in a given term divided by the total number of faculty members that had taught during the *prior* term.[iv] As noted earlier, at the outset of this study most of the college's instructors were hired on a part-time casual basis. Even after full-time faculty members were hired, they were employed on one year, three year, or five year contracts.

Thus, turnover was not constrained by the usual legal and contractual factors that characterize many universities.

Other than occasional relocation of spouses and the like, turnover in this context was generally a result of one of two causes. First, it could result from poor performance, in which case employment was not continued at the end of the term for a part-time instructor or at the end of the contract for a full-time instructor. Second, it could result from dissatisfaction or a sense of lack of fit on the part of the instructor.

Sample

The unit of analysis in the quantitative phase of this study is a course and its composite class sections. We employed data describing student perceptions of course sections taught between the third term in the 2000-01 academic year (i.e., January term, 2001) and the first term in the 2007-08 academic year (i.e., June term, 2007) for a total of 21 trimesters. Table 1 includes the sample characteristics for GSB students and instructors. During the period of the study, 1,739 course sections were taught by 233 different instructors to 40,686 students. There were 33,896 student questionnaires returned and included in the analyses.[v] Table 1 indicates that the student response rate across class sections was greater than 83%, meeting the requirements for this type of research (Lyon & Hendry, 2002).

Insert Table 1 about here

Data Analysis

Data analysis focuses on several issues. First we describe the design of the faculty evaluation system through reference to qualitative data drawn from the author's notes and GSB documents. Second, we examine the extent to which changes in student perceptions of Instructor Effectiveness were significant over time using *t*-tests. Finally, we assessed patterns of change in faculty turnover over the seven-year period.

Limitations

The assumption behind these analyses is that the system of the faculty performance assessment system was a central, though not the sole feature that contributed to these changes. As a non-experimental study, however, we acknowledge two primary threats to validity. The first is attributing the pattern of results to an unmeasured variable. In this study many changes were occurring simultaneously, so it is difficult to pin down the relative impact of different management tools (e.g., training, evaluation, reward). We view this threat as a cautionary admonition against drawing unwarranted interpretations. The second threat is failing to account for contextual factors that could have impacted the results. In this case, we acknowledge the rather unique context in which the case study was conducted and draw interpretations with this in mind (Yin, 2008).

Design of a Faculty Evaluation System at the Graduate School of Business

Although faculty evaluation represents one potentially effective approach to improving the quality of teaching in universities it has generally failed to achieve its potential (Bridges, 1986). Throughout the world, the implementation of faculty evaluation is constrained by labor laws, union contracts, institutional policies, capabilities of supervisory personnel, and norms of academic freedom and classroom privacy of the classroom (Aleomoni, 1999; Arreola, 1984; Bridges, 1986; Paulsen & Feldman, 1995; Rifkin, 1995). In its current state, faculty evaluation is often viewed by teachers and administrators as a legally-required but 'ritualistic exercise' (Meyer, Scott & Deal, 1983) with little or no instrumental value related to either monitoring or improving the quality of teaching in universities. Assessments of teacher evaluation confirm its negligible impact on teaching and learning, despite the inordinate amount of time and money invested in lengthy and detailed procedures (Bridges, 1986; Seldin & Angelo, 1997).

Despite this sobering assessment, features of the GSB's context seemed to suggest that data-based decision-making focused on teaching performance could enable a rapid improvement in the college. Given the employment conditions specified earlier, there was the possibility of using a rapid strategy of 'changing the people' rather than a tortuous strategy of getting the people to change. Given the ambitious vision and the highly traditional context in which the GSB was situated, there would be time enough for the latter as well. It was therefore decided to explicitly link faculty evaluation to instructor selection, and subsequently to reward and training.

Faculty Performance Assessment System

An assumption of virtually all quality improvement approaches is that the implementation of individual improvement approaches (e.g., faculty evaluation) must occur in the context of *systemic change* (Senge, 1990). Stated differently, no single improvement policy or practice will bring about significant, lasting change when implemented in isolation. With this in mind, a performance assessment system was designed with six key elements:

- Meaningful vision of learning;
- Course evaluation;
- Interim course evaluation;
- Use of a performance assessment rubric;
- Performance feedback;
- Reward for performance.

Meaningful vision of learning. Terrence Deal has observed that "culture is what keeps the herd moving in roughly the same direction" (personal communication, 1986). The quality improvement effort at GSB was clearly asking 'the herd' to change its direction. Given the nature of university instructors, rewards and punishments were unlikely to succeed in the

absence of a meaningful vision of learning. Therefore, all conversations about performance and reward were couched in terms of a vision of learning that centered on students.

The vision of learning was to produce graduates who were capable of using global and local knowledge to solve problems in organizations and the society. In order to make this vision more tangible, senior faculty actively encouraged the use of approaches to teaching that would foster action-directed learning among students. Training was provided to support the use of problem-based learning, video-enriched cases, team exhibitions, cooperative learning, cases, and role plays. The vision of problem-focused, practice-oriented, active learning became the continuous focus of GSB faculty deliberations and activities.

Underlying this change was an implicit refocusing of the college and its priorities from faculty to students. In Thai, the colloquial word used for students is *noo*, which translates to mean 'mouse'. This suggests the relative status of students in the educational hierarchy. It was an underlying, though explicit, objective to reorient management systems of the college, not only teaching and learning, around the needs of students. This was perhaps an even more radical change.

Course evaluation questionnaire. As indicated above the GSB developed its own CEQ (see Figure 1). Reliability and validity of the CEQ were relevant issues, especially since the instrument was used not only for diagnostic feedback, but also for decision-making. After 21 terms of use, data analysis indicates that the whole scale has an internal consistency of .964 (Cronbach's alpha).

As an initial attempt to check the validity of data obtained from the scale, student feedback was compared to data obtained from alternate sources. For example, data collected from observations of eight selected instructors was compared with data on several items. The instructor sample consisted of four highly-rated and four low-rated instructors where the

ratings were consistent across at least three terms. Three 20 minute observations were made in each of these instructor's classes. The observations focused on only four items:

- "encourages students to learn from each other"
- "actively involves students in learning"
- "makes clearly organized presentations" and
- "clear communication in English".

Available *documents* were also used to check the results for the same instructors on the selected items. These items included: quality of class handouts, the nature of examinations (e.g., whether they tested for understanding), and adequacy of course outline. While this was not a comprehensive validation study, the trend of the results of these analyses suggested that the questionnaire results were valid. Subsequent use of the instrument entailed periodic checks of data against alternative sources.

We earlier described the procedures for administration of the CEQ. Research has shown that non-systematic administration can result in bias and threaten the validity of student ratings (Aleamoni, 1999; Scriven, 1995). These issues were treated seriously by the academic administration at GSB for the very reason that the results would be used for personnel decision-making (Scriven, 1995).

Anecdotally, we note that in the Thai context this was actually more complex than might be imagined. Much like the students (i.e., *noo*), the academic staff members were 'junior' in age and status to the faculty. In Thai culture this meant that they were inherently reluctant to ask the instructors to leave the room during administration of the CEQs. This occurred initially in spite of explicit verbal and written instructions communicated to all parties. The management team, therefore, counted it as a 'quick win' (Kotter, 1996) on the day that a junior member of the academic staff directly asked the Academic Director himself/herself to leave the classroom when s/he stayed at the teaching table hile the CEQs

were being distributed. The Academic Director subsequently shared this story with other academic staff to confirm that 'it really was alright' for them to ask the faculty to leave the classroom.

Interim course evaluation. The GSB management team was intent on using the CEQ results for summative evaluations of performance that could be used in decision-making on the selection of instructors. However, equal in priority was ensuring that useful feedback would reach instructors in a timely fashion. With this in mind, a short Interim Course Evaluation questionnaire was designed for the purpose of providing formative student feedback to instructors during the term.

The Interim Course Evaluation consisted of three Likert-type questions, rated on a 1-5 scale using Poor to Excellent response categories, and two open-end questions. The Likert questions asked students to rate the instructor on: 1) The instructor's preparation for class; 2) The instructor's ability to actively involve students in learning; 3) Your overall rating of the course so far. The open-end questions were: 1) What I like most about this class is. . . ; 2) This class will be better if. . . .

These Interim Course Evaluations were collected by academic staff midway through the course. Copies were made of the interim evaluations from these classes and sent to the Academic Director of the college. Program Chairs would have two days to review the results from their programs before sending the original evaluation documents to faculty members. Faculty received the documents within five days of their having been completed by students. This ensured that instructors had sufficient time to make mid-course corrections during the term and program administrators had early warning so they could help solve small problems before they became larger.

Performance assessment rubric. Another key element in the GSB's performance assessment system was the design and use of an analytical rubric for

assessment of faculty performance of full-time instructors. This simple assessment rubric (see http/www.xxxx.html) incorporated the three common domains of university instructor performance: teaching, research, service. With respect to teaching, results on the CEQ and improvements in the teaching process were included as distinct categories on the rubric. The rubric also incorporated a system whereby instructors could assign weights to criteria within certain pre-defined limits.

In practice, the faculty member completed the rubric before submitting it with a short narrative and relevant supporting documents. These would be reviewed by the Academic Director who would pencil in his/her own initial assessment. The final assessment would be made following a meeting between the Academic Director and the faculty member. The results on the rubric were used to determine the extent of reward, as described below.

Performance feedback. Feedback on performance was considered an essential part of the evaluation system. That is, many universities collect data without making available to faculty in timely and useful ways. It was, therefore, a priority to ensure that faculty received their feedback from students as rapidly as possible following completion of the interim and final course evaluations. A data collection and management system was designed to ensure that instructors received their summarized feedback within five days of completion or the Interim Evaluations and 10 days for the end of term evaluations.

In terms of direct feedback, at the individual level, program managers would discuss teaching results with their relevant faculty members prior to the start of each succeeding term. These discussions could be quite informal and usually took place while discussing course plans for the following term. When results fell below expectations, joint problem-solving usually occurred.

Full-time faculty also reported directly to the Academic Director of the GSB. Unless there was a specific problem, performance feedback from the Academic Director occurred twice a year in concert with administration of the college's reward system. CEQ results were incorporated into the performance rubric and discussed at individual faculty conference sessions (i.e., twice a year in September and March).

In addition to these individual performance reviews, other steps were taken to refocus the culture of the college on the quality of teaching and learning of students. For example, at the pre-term faculty meeting held each trimester, the Academic Director would present and discuss the GSB faculty profile from the previous term. Sometimes this would lead to suggestions for the faculty as a group to consider or even to changes in policies (e.g., with respect to assessment). These were key opportunities to fine-tune the message of what was expected of the faculty whether it concerned active learning, quality of syllabi, modes of assessment or punctuality.

Reward. One common cause of failure of performance evaluation in higher education is the absence of meaningful rewards. The GSB system incorporated several types of rewards and spaced them at regular intervals. The principles underlying the construction of the reward system was to reward on a mastery not a normative model, to reward many for high performance rather than one or two, to address intrinsic as well as extrinsic motivators, and to ensure that rewards were distributed with sufficient frequency that faculty would link the rewards to their performance.

First, the annual raise at the GSB for full-time instructors could range from as low as 1% to as high as 10%. Although this ran counter to common practice in Thai universities, it was a priority of the management team to ensure that top performing faculty got substantially higher performance awards than lower performing faculty.

Second, at the mid-year period, an interim evaluation was conducted using the same performance rubric. Assessment was based on the first half-year's performance. It would result in a one-time bonus of between USD $200 and USD $2,500. Again, this approach was based on principles of pay for performance, making the degree of award meaningful, and timing the feedback close to the period of performance.

Finally, in order to create a broader impact on the faculty culture, a system of teaching excellence awards was initiated for distribution at the pre-term faculty meeting three times per year. Five individual teaching awards and one team teaching award were given out, based largely but not wholly upon the CEQ results. The reward was distributed in a ceremony that involved handing a certificate of teaching excellence and a check for $250 USD to each awardee individually. Each recipient would then 'tell a story' that sought to communicate what s/he had done in his/her class that term which had created student success. The names of recipients were then posted on visibly in the college for posterity.

As the reader would expect, during the first few terms this celebratory ritual was greeted with no small degree of cynicism among the faculty, especially those who did not gain awards. However, over time, as the culture of the GSB changed, the cynics were won over. Indeed, some of the early critics even visited the classrooms of other faculty to 'see what they did' in later years.

The intrinsic nature of these awards should not be underestimated. One Associate Professor from another university who taught part-time at the GSB hung around waiting to speak to the Academic Director one night after receiving his award. With a breaking voice he said, "You know, I am an Associate Dean at my university and I've taught for 15 years, but nobody at my own university has ever taken the time to comment on or acknowledge the quality of my teaching. Here, I am just a part-time

instructor, but I can gain this kind of feedback and recognition. I can't tell you how much this means to me. Of course the check is welcome, but it really the recognition in front of the other faculty whom I really respect as fine teachers that makes this meaningful to me. If there's anything I can do to help the college in the future, please let me know" (personal communication, June 12, 2003).

This description of the quality improvement tools has outlined the main features of the *systemic change* that the management team sought to achieve. While anecdotal description is useful, it is subject to bias and necessarily offers an incomplete picture of 'what happened'. Therefore the next section focuses on quantifiable data that describe the results of implementation.

Implementation Results

The descriptive analyses included in Table 2 and Figure 2 show the change in student perceptions of Instructor Effectiveness over the seven-year period of the study. The data portray a pattern of rapid improvement in Instructor Effectiveness during the first four terms of implementation. The mean Instructor Effectiveness rose from a baseline of 3.84 to a level of 4.20 in term 6 and remained at or above this level for the remaining 14 terms.[vi] This suggests that the GSB was able to sustain stable results at this very high level term-by-term for a period of five years (see Figure 2).

Next we wished to determine if these seemingly substantial changes were statistically significant. As faculty members at universities throughout the world have observed, it is often difficult to determine whether the difference between a 3.81 and a 3.95 is 'really meaningful' or not. We used a *t*-test (see Table 2) to examine Instructor Effectiveness in the baseline term (term 0) and the final term of the study (term 20). This found that the difference (.44) was statistically significant ($t = -7.01$, $p < .01$). We note that additional *t*-tests

confirmed the significance of change from the baseline level beginning with the third term of implementation and for each subsequent year. We note that in a separate study, mixed models (Heck, Thomas, & Tabata, 2010), a more sophisticated test of change over time, confirmed the findings from the *t*-tests used in this study (Author, in press). Moreover, this test further established a pattern of linear growth in scores that was statistically significant year-on-year over time.

Thus, we conclude that improvement in Instructor Effectiveness over time was both substantial in terms of the scale of the change and statistically significant. The improvement of .44 in mean scores between term 0 and term 20 represented a positive gain of a full standard deviation in the instructor profile. Moreover, the size of these differences as perceived by students was also meaningful in the sense that it reflected trends in other data collected in the college. For example, both the number and quality of student applications rose annually following the improvement in faculty performance. Results on the GSB's comprehensive exam also demonstrated a higher rate of passage as well. Thus, we suggest that improvement in instructor effectiveness can be characterized as substantial, significant, and meaningful.

Drawing on the quality literature, however, we had a second goal which was to reduce the variance in Instructor Effectiveness ratings. That is, use of the performance assessment tools also sought to reduce variation in instructor effectiveness across different courses and course sections (i.e., many courses were taught in multiple sections by multiple instructors). Reducing variance would imply greater consistency in the education offered to students across classes. The goal of reducing variance in instructor performance was monitored on a term-by-term basis. This broad goal was, in fact, translated into three operational objectives:

1. To increase the percentage of instructors falling in the upper bands of the distribution;
2. To reduce the percentage of courses falling between 3.00-3.50;
3. To eliminate the occurrence of courses falling below 3.00.

The data in Table 2 and Figure 2 offer insight into trends related to these objectives. During the baseline term in early 2001, the standard deviation of Instructor Effectiveness ratings was .54. This indicates that 68% of Instructor Effectiveness ratings fell between 3.30 and 4.38, a very large range. Within two years, however, the SD on these two indicators had been reduced by a full .20, indicating that 68% of Instructor Effectiveness ratings fell between 3.74 and 4.42. Note that the initial improvement was obtained largely by reducing the number of courses receiving poor ratings (also see Figure 2). In the final year of the study, this instructor profile had further improved, whereby one SD of course ratings fell between 3.89 and 4.63. The lower end of the first standard deviation had risen above the mean level of performance during the baseline term.

Experience with this scale in the college suggested that courses where the Instructor Effectiveness rating fell below 3.50 (about 1 SD below the mean) were experiencing a level of problems that warranted problem-solving between the instructor and his/her supervisor. Over time, part-time instructors who earned a rating below 3.25 were seldom invited back to teach. Of course, these decisions were not made in isolation based only on course evaluation data from a single term. Prior results were taken into account as well as extenuating circumstances related to a particular class. Coaching was offered where it was needed.

Figure 2 portrays the clarity of this trend graphically. The min/max analysis highlights the reduction in the percentage of courses that received poor ratings from students. First, consistent with the variance analysis, we note that the *full distribution* of course ratings (i.e., difference between min/max) in a given term was reduced from 3.50 in the baseline

term to an average of about 2.20 in subsequent terms. While courses rated below 3.00 were not fully eliminated, they did reach the status of *rare species* over time.

The final analyses conducted for this study concerned faculty turnover. In its initial years, the GSB relied almost entirely upon part-time instructors. This was an intentional strategy employed by the founding Director to keep fixed costs low during the start-up phase of the college. The quality audit in 2000, however, made it clear that the college required a complement of full-time faculty to support sustainable academic development of the college.

It was, however, a goal of the GSB management team to maintain a balance between part-time and full-time faculty members. While the optimal balance was subject to debate, financial and human resource constraints, as well as concerns over teaching quality moderated efforts to increase the percentage of full-time faculty beyond a certain level. Moreover, the GSB management was aware that finding full-time instructors who could *both* conduct high quality research *and* teach to its level of expectation was difficult. Even as research began to assume greater importance from 2003 onwards, the management team was reluctant to compromise on the quality teaching. Too much had been invested in reaching this point.

GSB managers believed fundamentally that part-time faculty had much to offer in teaching how to manage organizations in the local environment. Moreover, the GSB covered a wide range of specialization programs and it could be difficult to find full-time instructors who were truly qualified to teach certain subjects (e.g., Enterprise Resource management, Logistics, Quality Management). Part-time instructors were needed to fill this gap.

Data in Table 2 indicate a gradual growth in the number of full-time instructors, stabilizing at about 15 in the latter years of the study. Over time, the percentage of courses taught by full-time and part-time faculty also stabilized at about 40% and 60% respectively. We should note, however, that the GSB's human resource strategy sought to create a close

relationship with the part-time faculty who taught in the college. Indeed, the GSB's curriculum structure made this essential. Core courses (five), Foundation Courses (about eight across different programs), and the GSB's PBL capstone track (seven modules) all relied on instructional teams for the implementation of common curriculum units. This approach required an extremely high level of faculty interdependence and meant that continuity in instructor teams was critical. Reliance on part-time faculty to assist in fulfilling this need entailed risk but the potential payoff was large. Thus, the GSB management formulated a human resource strategy designed to transform the part-time faculty into 'long-time faculty'.

This strategy entailed offering premium pay relative to the local context, more or less guaranteeing regular employment term-by-term, providing optimal teaching environment (e.g., small class size, excellent facilities), issuing adjunct appointments with name cards, and offering eligibility for various benefits normally reserved for full-time faculty (e.g., conference travel). In return, the college expected the 'long-time faculty' to attend team planning and grading meetings, training workshops, and other GSB events. In end, students could seldom distinguish between which faculty members were 'long-time' or full-time.

The success of this strategy is reflected not only in the overall Instructor Effectiveness results but also in data on faculty turnover (see Figure 3). During the first four terms of implementation of the performance assessment system, the rate of faculty turnover was shockingly high, averaging almost 25% per term. About 70% of the instructors who had been teaching during the baseline term were no longer teaching at the GSB in term four. This was due largely to the use of the performance assessment system both for instructor selection and training and development.

Figure 3 shows that by term five, the turnover rate was sharply reduced to a level below 10%. Except for a moderately sharp rise during a period of political instability in 2004,

it remained at this acceptable level for the remainder of the study period. We characterize 10% as an 'acceptable level' in light of the large number of instructors in the total pool and the goal of maintaining high levels of teaching performance.

Discussion

> "You can't mandate what matters to people, but what you mandate matters, (McLaughlin, 1990, p. 14)

Milbrey McLaughlin's quote offers useful insight into this case study. The quality implementation effort sought to bring about systemic change in the GSB. Yet, success depended upon obtaining the cooperation of the faculty (and staff), and motivating them to take ownership for making the new vision a reality. The management team initially 'mandated' many features of the new quality systems, knowing all the while that this would raise the level of faculty resistance.

Part of the quality effort entailed reorganization into a more centralized matrix organizational structure. Previously the Director had been at the center of a wheel structure with the Program Chairs at the end of the spokes. These independently contracted managers then hired whomever they wished to teach, typically their 'friends' from other universities. Given the substantial fee for part-time teaching, this was a form of interpersonal currency. These part-time instructors taught whatever it was that they taught at their own university. This meant that programs were loosely coupled in terms of curriculum, Program Chairs had minimal interaction with one another, and faculty seldom knew who else was teaching, even in the same program.

Not surprisingly, however, the use of data-based decision-making in instructor selection ran counter to the 'know who' culture of the college. Under the new structure, the Academic Director had the final decision on instructor selection and met personally with all new instructors prior to their being hired, even for one term. Course evaluation data became a

key tool for making these decisions. When the new evaluation and selection systems came into effect and turnover accelerated, there was a predictable reaction from 'friends' who questioned both the loyalty and authority of the Program Chairs (e.g., "I thought that you were the one who made these decisions. You mean you don't?"). Ultimately, five of the seven Program Chairs either resigned or were removed from their positions due to their unwillingness to accept the new systems.

One memorable event occurred the day that the Academic Director refused to sign off on the reappointment of a part-time faculty member who happened to be a close friend of the GSB Director. This instructor, a former Dean from another local university, had received the lowest ratings in the college as well as many highly negative comments from students for two consecutive terms. A politely waged tug of war, in Thai, ensued between the Director and the Academic Director. In the end, the decision not to reappoint the Director's friend stood, but in Redmond's (1994) words this and other similar incidents left " rents in the social fabric and frayed edges of tender feelings" (p. 14).

Little did Karl Weick (1976) imagine that 'loose coupling' could be taken to these limits. Nonetheless, hidden in this chaotic situation was opportunity. The loosely coupled nature of program structures and personal relationships reduced the likelihood of organized resistance. As a result, resistance to these changes took the form of many small arrows being shot continuously at the new management team for the first year, but with relatively little effect.

Despite these bumps in the road towards change, the results presented in this study confirmed a positive trend in development of teaching quality at the GSB. We attribute this result, at least in part, to the systemic human resource approach built around a system of performance evaluation. Analysis of this substantial longitudinal dataset demonstrates the largely untapped potential that performance assessment can have when implemented under

"the right conditions." In this section we wish to briefly comment on the nature of those conditions at the GSB and other institutions of higher education.

First, we suggest that quality improvements were positively impacted by the *systemic* approach that the College took towards organizational change (Senge, 1990). The new evaluation tools were implemented *systemically* and *simultaneously* with a clearly focused vision, new policies on reward and instructor selection, training and support for the use of new teaching methods, and regular performance feedback. We believe that the systemic implementation of change provided the pressure and support necessary to overcome resistance to change and equip faculty with the tools needed to make the new vision a reality.

Other organizational conditions that supported the successful implementation at the GSB have been mentioned either explicitly or implicitly in the discussion of results. Unequivocal support from the GSB Director gave the management team under the leadership of the Academic Director freedom to innovate and execute as it deemed suitable. Freedom from the constricting labor regulations of the parent university and lack of a union contract enabled to the GSB to use teaching performance as the main criterion for renewal of contracts both for full-time and part-time faculty.

Although the GSB operated in a highly competitive environment, few of its competitors were notable for teaching quality and none for research capacity. This created an opportunity. It meant that the GSB could, for the short to medium term, focus on a 'unitary goal', quality of teaching and learning. This was a luxury that comparable universities in more developed environments could not afford. Given a single goal and a set of potentially powerful management tools, the GSB management was able to overcome some of the latent inefficiencies that characterize universities as loosely coupled organizations (Weick, 1976).

The results were quite dramatic. In the third year of implementation, to the shock of the local business school community, the GSB won both local and international business challenge competitions. As word of the quality of teaching and learning at the GSB spread in the local market and among employers, application numbers increased steadily. Even though GSB operated a large Master degree program, admitting about 375 students per year, this enabled the college to improve its selectivity from 98% admissions in the year prior to the quality implementation to about 40% seven years later.

The reliance on quantitative data to tell this story is both a strength and weakness of this report. On the one hand it offers a more objective picture of 'what happened' over time than a story that relies heavily on the highly subjective narrative offered by a key participant in the change effort. On the other hand, the focus on quantitative data has not done justice to the heart of the change effort as embodied in the increased motivation and work ethic of faculty and students. While space does not permit elaboration, we close with the conclusion that the implementation of this quality did bring about change in the culture of the GSB over a seven year period.

Epilogue

The epilogue to this story borrows a page from Kotter's (1996) book, *Leading Change,* in which he asserts that a common error in change implementation lies in "declaring victory too soon." After more than seven years of steady progress, one would think that it would be possible to 'declare victory.' However, events took an unexpected turn in January 2007, when the current GSB Director,[vii] a staunch supporter of the quality improvement effort, retired in the middle of the academic year. The University's President, who also served as Chairperson of the Board of Trustees of the GSB, promptly appointed himself as Acting Director of the College.

Within a matter of months, he determined that the systems used in the GSB looked too dissimilar from those used by the University proper. At a meeting of GSB faculty and staff in the late spring, the President proclaimed that the GSB was "not being properly managed." This pronouncement signaled the beginning of a new change, a 'harmonization campaign' to bring GSB policies, systems, and practices in line those of the parent university. Our interpretation was that that the 'center of innovation' had become too successful in achieving its mission. The GSB has transformed to such an extent that its culture was simply too different from the culture of the larger university to be considered 'legitimate' despite its positive results (Meyer, Scott & Deal, 1983). It turned out that the politico-cultural support of the former Director, one of the most influential administrators in the University, had been a 'hidden success factor.' He had created the necessary 'space' for the cultural transformation to evolve. Once his 'protection' was removed, the dominant corporate culture swamped the 'island of competence' at the GSB.

At the end of the first term in the 2007 academic year (i.e., the last term for which data was included in this study), the Academic Director and several key members of the management team resigned. A new management team was selected largely from outside the GSB. The new Director and his team were well-known Thai administrators, trusted to use well-accepted methods in running the college. Although the quality tools used by the previous GSB management team nominally remained in place, they were quickly denuded of their impact and transformed into ritualistic exercises. For the most part, junior staff continued to move paper through the prescribed paths at the appointed times, but performance reviews were no longer held with individual faculty, data were no longer discussed with the faculty as a whole, and rewards were reduced in size and revised to reflect a more 'typical' range. Harmonization quickly gained the upper hand.

As it is not our purpose to extend this story beyond the suitable length of an epilogue, we leave the final word in the voices of GSB students. After *eight consecutive years* of maintaining enrollment at a level of 360 to 400 new students per year, nine months after the launch of the harmonization campaign, the succeeding intake of students dropped to 180. Word of mouth from current students travelled fast and prospective applicants voted with their feet. Even so, the 'new change' was considered a success by the university's senior administration as the GSB began to take on the 'appearance' of a Thai university. Legitimacy was being restored.

References

Aleamoni, L. (1999). Student ratings myths versus research facts from 1924 to 1998. *Journal of Personnel Evaluation in Education, 13*(2), 153-166.

Altbach, P., & Umakoshi, T. (2004). *Asian universities: Historical perspectives and contemporary challenges*. Baltimore, Maryland: John Hopkins University Press.

Arreola, R. A. (1984) Evaluation of faculty performance: Key issues. In P. Seldin (Ed.), *Changing practices in faculty evaluation: A critical assessment and recommendations for improvement* (pp. 79-85). San Francisco: Jossey-Bass.

Astin, A. (1999). Student involvement: A developmental theory for higher education. *Journal of College Student Development, 40*(5), 518-529.

Bridges, E. (1986). *The Incompetent Teacher: The challenge and the response*. Philadelphia: Falmer Press.

Creswell, J. (2007). *Designing and conducting mixed methods research.* Thousand Oaks, CA: Sage.

GSB. (2000). *Academic quality audit report*. Unpublished report presented to the Board of Trustees. Bangkok, Thailand: School of Business.

Greenwald, A. G., & Gilmore, G. (1997). Grading leniency is a removable contaminant of student ratings. *American Psychologist, 52*(11), 1209-1217.

Hallinger, P., & Bridges, E. (2007). *Preparing managers for action.* Dordrecht, Netherlands: Springer.

Harrison, P. D., Ryan, J. M., & Moore, P. (1996), College students' self-insight and common implicit theories of rating of teaching effectiveness. *Journal of Educational Psychology, 88*(4), 775-782.

Heck, R. H., Thomas, S. L., & Tabata, L. N. (2010). *Multilevel and longitudinal modeling with IBM SPSS.* New York: Routledge Academic.

Holmes, H. & Tangtongtavy, S. (1996). *Working with Thais: A guide to managing in Thailand*. Bangkok: White Lotus.

Jones, D.A. (2009, Dec. 18). Teacher evaluation. *Chronicle of Higher Education*. Retrieved online July 30, 2010 from http://chronicle.com/blogPost/Teacher-Evaluations/19400/.

Kotter, J. (1996). *Leading change.* Boston, MA: Harvard Business School Press.

Lawler, J. (2008). *Strategic human resource management*. Thousand Oaks, CA: Sage.

Lyon, P., & Hendry, G. (2002). The use of the course experience questionnaire as monitoring evaluation tool in a problem-based medical programme. *Assessment & Evaluation in Higher Education, 27*(4), 339-352.

Marsh, H. W. (1981). The use of path analysis to estimate teacher and course effects on student ratings of instrument effectiveness. *Applied Psychological Measurement, 6,* 47-60.

McKeachie, W. J. (1997). Student ratings: The validity of use. *American Psychologist, 52(*11), 1218-1225.

McLaughlin, M. (1990). The Rand change agent study revisited. *Educational Researcher, 5,* 11-16.

Meyer, J., Scott, R., & Deal, T. (1983). Institutional and technical sources of organizational structure: Explaining the structure of educational organizations. In J. Meyer and W.R. Scott (Eds.) *Organizational environments: Rituals and rationality* (pp. 45-67), Beverly Hills: Sage.

Park, S. H., & Luo, Y. (2001). Guanxi and organizational dynamics: organizational networking in Chinese firms. *Strategic Management Journal, 22*(5), 455-477.

Paulsen, M. B., & Feldman, K. A. (1995). *Taking teaching seriously: Meeting the challenge of instructional improvement*. (ERIC Document Reproduction Service No. ED396615).

Redmond, M. (1994, Jan. 9). The unselfishness of not being there. *The Nation*.

Rifkin, T. (1995). *The status and scope of faculty evaluation*. (ERIC Reproduction Service No. ED385315)

Scriven, M. (1988). The validity of student ratings. *Instructional Evaluation, 92*(2), 5-18.

Scriven, M. (1995). *Student ratings offer useful input to teacher evaluations*. (ERIC Reproduction Service No. ED39824)

Seldin, P., & Angelo, T. A. (1997). *Assessing and evaluating faculty: When will we ever learn? (To use what we know).* Proceedings of the AAHE 1997 Conference on Assessment and Quality Assessing Impact: Evidence and Action.

Stockham, S. L., & Amann, J. F. (1994). Facilitated student feedback to improve teaching and learning. *Journal of Veterinary Medicine, 21*(2), 51-55.

Walker, A., Bridges, E., & Chan, B. (1996). Wisdom gained, wisdom given: Instituting PBL in a Chinese culture. *Journal of Educational Administration, 34*(5), 12-31.

Weick, K. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, *21*, 1-19.

Yin, R. (2008). *Case study research: Design and methods*. Thousand Oaks, CA: Sage.

GSB

Course Evaluation Form

1. Rate the instructor's ability to communicate clearly in English.
   (1) Poor (2) Not very good (3) Average (4) Very good (5) Excellent
2. Rate the helpfulness of the instructor outside of class.
   (1) Poor (2) Not very Good (3) Average (4) Very good (5) Excellent
3. Rate the instructor's grading compared with other courses taken at GSB.
   (1) Easier (2) Somewhat easier (3) Average (4) Harder (5) Much harder
4. Give an overall rating to this course.
   (1) Poor (2) Not very good (3) Average (4) Very good (5) Excellent
5. Rate the information provided by the instructor in the course outline.
   (1) Poor (2) Not very good (3) Average (4) Very good (5) Excellent
6. Rate the instructor's knowledge of the subject taught in this class.
   (1) Poor (2) Not very good (3) Average (4) Very good (5) Excellent
7. Rate the instructor's preparation for classes.
   (1) Poor (2) Not very good (3) Average (4) Very good (5) Excellent
8. Rate the instructor's punctuality – starting and ending classes on time.
   (1) Poor (2) Not very good (3) Average (4) Very good (5) Excellent
9. Rate the instructor's feedback on assignments.
   (1) Poor (2) Not very good (3) Average (4) Very good (5) Excellent
10. Rate the organization of presentations and exercises by the instructor.
    (1) Poor (2) Not very good (3) Average (4) Very good (5) Excellent
11. Rate the level of reading materials used in the class.
    (1) Very Easy (2) Easy (3) Suitable (4) Difficult (5) Very difficult
12. Rate the instructor's patience and clarity in responding to student questions.
    (1) Poor (2) Not very good (3) Average (4) Very good (5) Excellent
13. Rate the effectiveness of handouts provided by the instructor.
    (1) Poor (2) Not very good (3) Average (4) Very good (5) Excellent
14. Rate the ability of the instructor to make the course content practical.
    (1) Poor (2) Not very good (3) Average (4) Very good (5) Excellent
15. Rate the instructor's ability to actively involve students in learning.
    (1) Poor (2) Not Very Good (3) Average (4) Very Good (5) Excellent
16. Rate how well tests and assignments assessed your understanding of topics.
    (1) Poor (2) Not Very Good (3) Average (4) Very Good (5) Excellent
17. Rate the instructor's effectiveness in encouraging students to learn from each other.
    (1) Poor (2) Not Very Good (3) Average (4) Very Good (5) Excellent

Please respond to the following questions in space provided on the reverse side.

1. The strongest features of this course were
2. The weakest features of this course were
3. My suggestions for improving this course are
4. My suggestions for the instructor to improve his/her teaching are

Figure 1. *Course Evaluation Questionnaire*

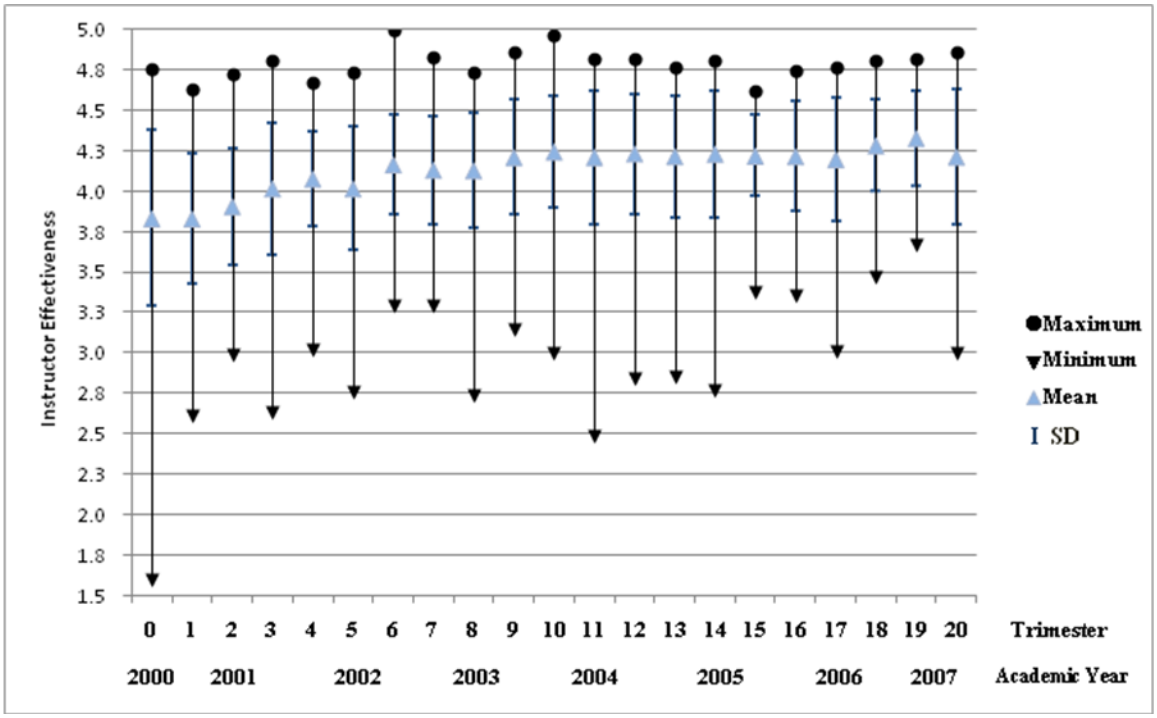

Figure 2. *Trend in Instructor Effectiveness, 2000-2007*

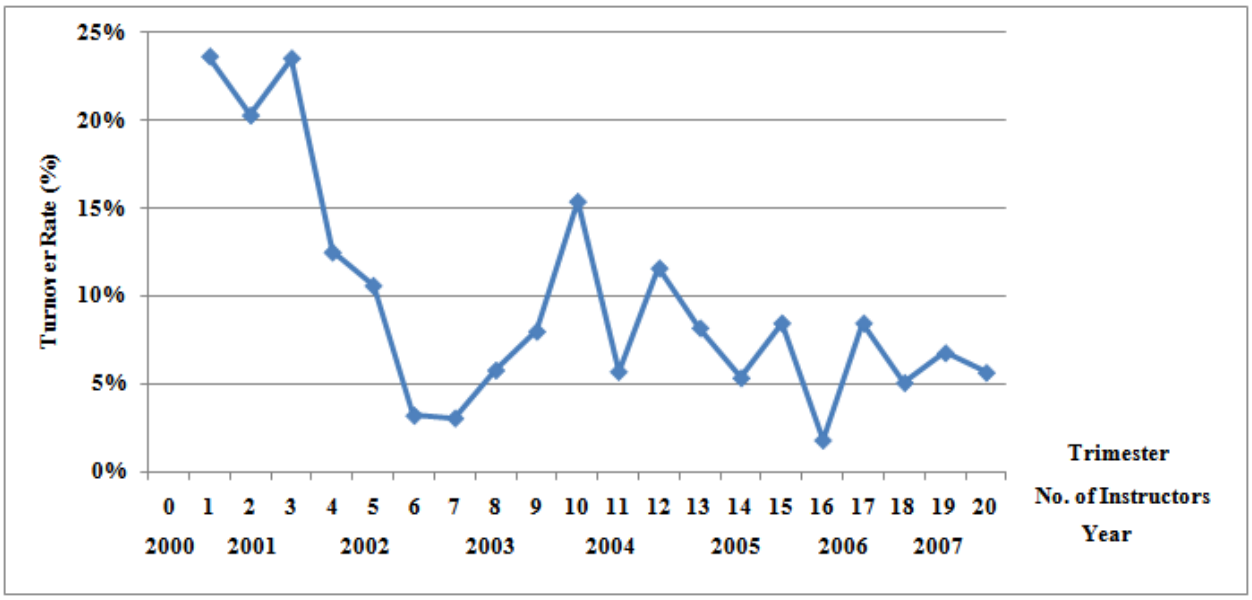

Figure 3. *Instructor Turnover Rate, Term-by Term, 2000-2007*

Table 1. *Summary of Course, Instructor, and Student Information: 2000-2007*

Students, Instructors and Classes	**Total**
Number of Course Sections	1,739
Total Number of Instructors	233
Average Students per Class Section	24.47
Total Number of Students in Courses	40,686
Total Returned Questionnaires	33,896
Response Rate	83%

Note:

a. The data for number of students and number of returned questionnaires in 2000 was absent.
b. The 'year' refers to the academic year. Data collection began in the third term of the 2000-01 academic year.

Dimension	N of items	alpha	2000 M (SD)	2001 M (SD)	2002 M (SD)	2003 M (SD)	2004 M (SD)	2005 M (SD)	2006 M (SD)	2007[c] M (SD)	Total M (SD)	t	Sig.
Mean Instructor Effectiveness	4	0.95	3.84(.54)	3.91(.39)	4.08(.34)	4.15(.35)	4.23(.38)	4.22(.35)	4.23(.34)	4.26(.37)	4.12(.40)	-7.01	**
Instructor Profile													
Instructors: Full/Part-time			8/46	12/87	14/49	13/66	14/54	13/44	17/43	14/42			
Courses Taught by:													
Full-time Instructors			12	88	72	87	71	82	86	58	556		
Part-time Instructors			64	223	151	196	172	146	138	93	1183		
Total Courses per Term			76	311	223	283	243	228	224	151	1739		

Table 2. *Descriptive Statistics for Change in Instructor Composition and Effectiveness: 200-2007*

Note: M = Mean; SD = Standard Deviation; n.s. = not significant; * = $p < .05$; ** = $p < .01$.
a. The statistics in the columns 2000 integrated the data of one trimester in 2000.
b. The statistics in the columns 2001 to 2006 integrated the data of three trimesters each year.
c. The statistics in the column of 2007 integrated data of two trimesters in 2007.

[i] In subsequent years a Thai language M.M. program serving about 600 students was opened at a separate campus, as well as a Ph.D. program on the campus of the international program.

[ii] It was anticipated that in future years government universities would be moved out of the traditional funding and governance structure. Therefore, the administration saw the opening of the GSB as an opportunity for the university to gain experience managing the institution outside the government system.

[iii] The author was initially a consultant hired to conduct the quality audit. Subsequently, s/he assumed the roles of Executive Director and Chief Academic Officer of the GSB during the seven-year period of this study.

[iv] For example, a faculty member who left the college due to her husband's job relocation was not counted as turnover for the purposes of this analysis which sought to highlight team solidarity.

[v] For the baseline term (i.e., January 2001), data on the number of students was missing from the CEQ data set.

[vi] Actually, the true baseline would have been the second semester of 2000 before any of the new quality initiatives had been put into place. However, the college was so disorganized at that point that the academic office could not even manage to distribute the printed CEQs to classes.

[vii] The GSB Director referred to here succeeded the founding Director who was forced to resign in 2003. The first Director supported the quality improvement effort up to a point, and then played a passive role, neither supporting nor directly opposing it. His successor gave unequivocal support to the implementation of the strategies described in this report.