

วารสารวิธีวิทยาการวิจัย

ปีที่ 10 ฉบับที่ 1 มกราคม - มิถุนายน 2540

JOURNAL OF
RESEARCH
METHODOLOGY

Volume 10, Number 1 January - June 1997

A BIENNIAL PUBLICATION OF
DEPARTMENT OF EDUCATIONAL RESEARCH
FACULTY OF EDUCATION
CHULALONGKORN UNIVERSITY

JOURNAL OF RESEARCH METHODOLOGY

วารสารวิธีวิทยาการวิจัย ปีที่ 10 ฉบับที่ 1 (มกราคม - มิถุนายน 2540)

JOURNAL OF RESEARCH METHODOLOGY (ISSN 1857-2933) is published semesterly (January -June, July -December) by Department of Educational Research, Faculty of Education, Chulalongkorn University, Phayathai Road, Bangkok 10330 Thailand, to carry original reports of studies and analyses in education, educational research methodology, educational statistics, educational measurement and evaluation.

Ordering Information

Member 2-year subscriber rate is 200 Baht; 4- year rate is 380 Baht. Single copies of back issues are 60 Bath. Discounts are available for quantity purchases. Send orders to JEM subscriptions, Department of Educational Research, Faculty of Education, Chulalongkorn University BKK 10330, Thailand, fax 218-2578 (local) 662- 215-3568 (international).

Change of Address

Claims for missing or undelivered issues will be considered only if received at the JEM Office within 6 months of the month of issue. Requests for change of address must be received at least 1 month before the publication date of the first issue to be affected by the request.

Instructions to contributors

Three fully blinded copies of the manuscript should be submitted for blind reviewing. The manuscript should be typed double-spaced (including quotations, footnotes, and references) on 8 1/2 x 11 in. paper, with ample margins, and should run between 10 and 15 pages in typed length. The author's name and affiliation should appear on a separate cover page, and only on this page, to ensure anonymity in the reviewing process. An English abstract of 100-150 words must be included on a separate page. Manuscripts are accepted for consideration with the understanding that they are original material and are not under consideration for publication elsewhere.

Editorial review usually takes 1-2 months. All figures must be camera-ready. Manuscripts not conforming to these specifications will be returned to the author for proper style change.

Editorial Correspondence

All editorial correspondence and manuscripts relating to the journal should be sent to Prof. Dr. Somwung Pitiyanuwat, Faculty of Education, Chulalongkorn University, BKK 10330. E-mail address:somwung@chulkn.car.chula.ac.th

Copyright and Permissions

@ 1997 by the Department of Educational Research. No written or oral permission is necessary to reproduce a table, a figure, or an excerpt of fewer than 500 words from this journal, or to make photocopies for classroom use. Authors are granted permission, without fee, to photocopy their own material. Copies must include a full and accurate bibliographic citation and the following credit line; "Copyright [year] by the Department of Educational Research; reproduced with permission from the publisher." Written permission must be obtained to reproduce or reprint material in circumstances other than those just described. Please direct requests for permission or for further information on policies and fees to the Department of Educational Research Office.

Advertising

JEM Office. Department of Educational Research, Faculty of Education, Chulalongkorn University, BKK 10330. Telephone: 218-2586, fax 218-2578 (local) 622-215-3568 (international) (rates and dates available on request). E-mail address:somwung@chulkn.car.chula.ac.th

Journal of Research Methodology

Editors

Somwung Pitiyanuwat, Ph.D.
Chulalongkorn University, Thailand

Teara Archwamety, Ph.D.
University of Nebraska at Kearney, U.S.A.

Associate Editor

Nonglak Wiratchai, Ph.D.

Chulalongkorn University, Thailand

Suwimon Wongwanich, Ph.D.

Chulalongkorn University, Thailand

Siridej Sujiva, Ph.D.

Chulalongkorn University, Thailand

Editorial Board

Utumporn Jamornmann, Ph.D.

Chulalongkorn University, Thailand

Sirichai Kanjanawasee, Ph.D.

Chulalongkorn University, Thailand

Amornwich Nakornthap, Ph.D.

Chulalongkorn University, Thailand

Uthai Dulayakasem, Ph.D.

Sodsri Foundation, Thailand

Pitr Thongchan, Ph.D.

Srinakharinwirot University (Prasarnmit),
Thailand

Sor Wasna Pravalpruk, Ph.D.

Srinakharinwirot University (Prasarnmit), Thailand

Arunsri Anantrasirichai, Ph.D.

Ministry of Education, Thailand

Sampharn Punprug, Ph.D.

Khonkhan University, Thailand

Utane Panyo, Ph.D.

Chiangmai University, Thailand

Jay Samuels, Ph.D.

University of Minnesota, U.S.A.

Jong Seung Lee, Ph.D.

Chungnam National University, Korea

Rebecca Barr, Ph.D.

University of Chicago, U.S.A

Ho Wah Kam, Ph.D.

SEAMEO Regional Language Institute, Singapore

Ray Derricott, Ph.D.

University of Liverpool. United Kingdom

Zsuzsa Matrai, Ph.D.

National Institute for Public Education, Hungary

Susan M. Brookhart, Ph.D

Duquesne University, U.S.A.

Managing Editor : Auyporn Ruenrakul

CONTENT

1

**Semantics, Psychometrics, and Assessment Reform :
A Close Look at "Authentic" Tests**

James S. Terwilliger

15

**The Validity and Reliability of Information from a Student Annotation
Form for Work Samples in Portfolios**

Susan M. Brookhart

37

**A Development of Multilevel Causal Analysis Technique :
an Application of HLM Program**

Nikom Nak-ai

67

**The Evaluation of the Cooperation between the Thai Government and
the Asian Institute of Technology Project to Produce Engineering and
Technology Personnel**

*Boonreang Kajornsin Samnao Kajornsin Pornthip Chaiyaso
Panurut Rattayapas Tasanee Sahawat Wantani Thaitiang*

84

**The Evaluation of Cooperation between the Thai Government
and the Asian Institute of Technology Pertaining to the Thai Government
Budgetary Support: Phase V (B.E. 2533-2537)**

*Somwung Pitiyanuwat Auyporn Ruengtrakul Wilaiwan Pitiyanuwat
Sopana Takaew Panurut Rattayapas Wantani Thaitiang
Tasanee Sahawat*

Semantics, Psychometrics, and Assessment Reform: A Close Look at "Authentic" Tests

James S. Terwilliger*

ABSTRACT

The purpose of this paper is to raise questions about the claims that are frequently made by advocates of assessment reform. In particular, the focus would be the work of Wiggins who first introduced the concept of "authentic" assessment and who was one of the most influential critics of traditional assessment approaches. The author agreed that assessment practices were in need of reform but believed that the rhetoric of the reformers was misleading and largely unsupported by data. The alternative assessment procedures should be adopted in combination with more traditional forms of assessment as evidence of the educational and psychometric value of such alternatives became available.

* University of Minnesota

Paper presented as part of an invited symposium on "Theory and Practice in Large-Scale Performance Assessment" at the annual of the National Council on Measurement in Education, New York City, April 9, 1996.

Introduction

It is obvious to even the most casual reader of the literature on educational assessment that the field is currently undergoing a fundamental and profound transformation. The traditional concepts and methodologies associated with assessment are being questioned by a variety of critics including school reform advocates, subject matter experts, cognitive theorists, and others. In general, advocates for change recommend that assessments of achievement should be designed to reflect more precisely complex “real-life” performances and problems than is possible with short-answer and choice-response questions that characterize the typical standardized test currently employed in large-scale testing programs.

The purpose of this paper is to raise questions about the claims that are frequently made by advocates of assessment reform. In particular, my focus will be the work of Wiggins (1989a, 1989b, 1993) who first introduced the concept of “authentic” assessment and who is one of the most influential critics of traditional assessment approaches.

I wish to make it clear at the outset that I do not oppose some of the ideas that have been put forth by Wiggins and others. In fact, I agree that assessment practices are in need of reform. However, I believe that, as is often the case, the rhetoric of the reformers is misleading and largely unsupported by data. I fear that there is a danger that perfectly useful and appropriate assessment methods will be discarded in a rush to adopt a variety of other techniques of unknown psychometric and educational quality. I believe that alternative assessment procedures should be adopted in **combination** with more traditional forms of assessment as evidence of the educational and psychometric value of such alternatives becomes available.

Word Magic and Assessment Reform

The **American Heritage Dictionary** gives the following definitions of “authentic”:

1. a. Worthy of trust, reliance, or belief: **authentic records**. b. Having an undisputed origin: genuine.
2. Law. Executed with due process of law: **an authentic deed**. Synonyms are listed as “real”, “genuine”, and “authoritative”. Obviously, terms like these have a decidedly positive connotation. Therefore, any object or product to which these terms are applied is likely to be viewed as more desirable or of better quality than objects or products which are not so described.

That is why these terms appear so frequently in commercials and advertisements in the popular media, e.g. statements such as “Coke is the **real** thing!”, “**Genuine** factory auto parts, and Authentic French cuisine. A bakery chain in the Twin Cities has recently introduced RENAISSANCE BREADS with the slogan “Authentic handmade breads in the European tradition.” How can one doubt the quality of such a product when the price is \$3.79 per loaf?

The appeal of the term “authentic” is obvious in its widespread adoption by a variety of educators since it was first introduced into the literature. A variety of books with titles such as **Authentic Assessment in Action** (1995) and **Assessment of Authentic Performance in School Mathematics** (1992) have recently been published. Lesh and Lamon, the editors of the latter book, provide the following interesting definition:

Stated simply, authentic mathematical activities are those that involve: (i) real mathematics, (ii) realistic situations, (iii) questions or issues that might actually occur in a real-life situation, and (iv) realistic tools and resources. (p.18)

There are two problems with this definition. First, as previously noted, words like “real”, “realistic”, and “real-life” are synonymous with the word “authentic”. Therefore, the definition is circular. Second, and more fundamental, what constitutes “realistic” or “real-life” situations and “realistic” tools and resources is frequently open to debate. What appears to be “realistic” to one individual often seems to be “unrealistic” to another. The only way to “objectively” define “real-life” questions and situations would be to construct a data base that could be employed in defining the likelihood that a student would encounter specific questions, problems, etc. in non-school settings. Lacking such a data base, test designers typically rely on individual (or perhaps team) judgments of what is “realistic”.

It seems that the word “authentic” has an almost mystical power. Phrases such as “authentic instruction”, “authentic performance”, and “authentic outcomes” are appearing with increasing frequency in the educational literature. (One can only speculate about forms of instruction, performance, and outcomes that might be labeled “inauthentic”) I note that one of the NCME training sessions for this convention is titled “Ensuring that Authentic Assessment is Fair”. It is reassuring that there is an increasing recognition that even “authentic” assessments pose the same problems of “fairness” as do more conventional assessment techniques.

Clearly, individuals who are not sophisticated with respect to issues associated with the design and analysis of assessment procedures are more likely to view positively any procedure which is labeled “authentic” regardless of any conceptual, technical, or

practical issues that many arise in the application of the procedure. Educational assessment is a complex process which is built upon a variety of assumptions about the purposes of education along with a set of data gathering procedures which need to be judged against a series of both practical and technical standards. The use of labels which impute special status to a specific set of data collection procedures only serves to obscure more fundamental assessment questions that must be addressed. Therefore, terms like “authentic”, “genuine”, and “real-life” should be reserved for advertising copy and be avoided in scholarly discussions of educational assessment.

Origins of “Authentic” Tests

The term “authentic” was first introduced in reference to tests by Wiggins (1989a) in an article in *Educational Leadership*, a journal for school administrators and general educators. (This is an audience unlikely to include many psychometrically trained readers.) Wiggins defined “authentic” tests in terms of complex performances or exhibitions in which a student completes a report or makes a public presentation following an extended period of work on an out-of-class assignment. Wiggins presents the example shown in Figure 1.

Fig.1. An Example of a Test of Performance

An Oral History Project for 9 th Graders

To the student:

You must complete an oral history based on interviews and written sources and then present your findings orally in class. The choice of subject matter is up to you. Some examples of possible topics include: your family, running a small business, substance abuse, a labor union, teenage parents, and recent immigrants.

Create three workable hypotheses based on your preliminary investigations and four questions you will ask to test out each hypothesis

Criteria for Evaluation of Oral History Project

To the teacher:

- Did student investigate three hypotheses?
- Did student describe at least one change over time?
- Did student demonstrate that he or she had done background research?
- Were the four people selected for the interviews appropriate sources?
- Did student prepare at least four questions in advance, related to each hypothesis?
- Were those questions leading or biased?
- Were follow-up questions asked where possible, based on answers?
- Did student not important differences between “fact” and “opinion” in answers?
- Did student use evidence to prove the ultimate best hypothesis?
- Did student exhibit organization in writing and presentation to class?

Note: This example is courtesy of Albin Moser, Hope High School, Providence, Rhode Island. To obtain a thorough account of a performance-based history course, including the lessons used and pitfalls encountered, write to Dave Kobrin, Brown University, Education Department, Providence, RI 02912.

From Wiggins, 1989 (p. 44)

Wiggins concludes his discussion of "authentic" tests as follows:

In sum, authentic tests have four basic characteristics in common. First, they are designed to be truly representative of performance in the field; only then are the problems of scoring reliability and logistics of testing considered. Second, far greater attention is paid to the teaching and learning of the criteria to be used in the assessment. Third, self-assessment plays a much greater role than in conventional testing. And fourth, the students are often expected to present their work and defend themselves publicly and orally to ensure that their apparent mastery is genuine. (p.45)

It is instructive to examine the example Wiggins gives in light of the four characteristics he claims all "authentic" tests share. First, what "field" is represented in the oral history project? Since the choice of topic is left to the student, the "field" must be history with special emphasis upon techniques employed by historians who employ "first-person" sources in their research. In fact, this approach is hardly "truly representative of performance in the field" if the "field" is more broadly defined as history since the great majority of historians rely upon written rather than "first-person" source in their work.

Second, it is not clear that the criteria for evaluation of the project were shared with the students. (In fact, the example strongly suggests that the criteria were for the teacher only.) Therefore, how could they have been "taught and learned"? Even if the criteria were shared with the students in advance, what exactly would a teacher expect them to learn from them? Most students would use the criteria as a checklist to make certain they had satisfied the teacher's demands prior to turning their projects in for evaluation by the teacher! Since many of these criteria are very specific to the particular project, they appear to have limited value as general learning outcomes.

As an aside, it is not clear from the criteria presented exactly how they are to be employed in evaluating the projects. Since most of the questions posed can readily be answered "yes" or "no", it would be possible to devise a simple checklist. Obviously, a more elaborate scoring system could also be designed but the example provides no clues if that is the case. As is the case in all examples he gives in his writings, Wiggins provides no data with respect to the reliability of scoring (or the amount of time devoted to scoring) the "authentic" tasks he recommends. The lack of supporting data is a reflection of the secondary role Wiggins gives to such issues in his reference to "scoring reliability and logistics of testing" in the above quote.

Third, with respect to self-assessment, there is no indication that this is involved in the oral history project.

Lastly, it is clear that the oral history project does involve an oral presentation to the class. Presumably, the oral presentation may also be followed by a discussion during which the presenter would have to answer questions and defend her/his work.

“Authenticity” and Validity

Several issues are highlighted through the detailed comparison of the oral history example with the criteria for “authentic” tests given by Wiggins. (The same example and criteria for “authenticity” are presented in a follow-up article (1989b) which appeared in the **Phi Delta Kappan**, another journal for a general readership in education.) First, it is not entirely clear what is meant by the phrase, “designed to be truly representative of performance in the field”. What exactly is the “field” in this example? Who decides what is “truly representative” of the field?. This raises questions regarding test **validity**. In his follow-up article Wiggins’ only reference to validity of “authentic” tests is the comment, “Far greater attention is paid throughout to the test’s ‘face’ and ‘ecological’ validity.” (1989b,p.712) However, in his more recent writings, Wiggins acknowledges that “face” validity is not highly regarded by measurement specialists and states,

Thus, although face validity should be considered, to focus only on it is to miss a more important point about the incentives to perform well that might be found to inhere in more authentic forms of assessment and might change the implications of the scores. (1993, p.244)

This statement seems to imply that interpretations (hence the validity) of results of “authentic” assessments is somehow enhanced by virtue of their being perceived as more “genuine” than conventional forms of assessment. Of course, this is one of the major arguments frequently advanced in support of “face” validity.

Messick (1994) has discussed at length issues associated with the validation of performance assessments. He cites the classic treatment of performance tests by Fitzpatrick and Morrison (1971) in which they state, “there is no absolute distinction between performance tests and other classes of tests” (p.238) Using this as a point of departure, Messick states,

Hence performance assessments must be evaluated by the same validity criteria, both evidential and consequential, as are other assessments. Indeed, such basic assessment issues as validity, reliability, comparability, and fairness need to be uniformly addressed for all assessments because they are not just measurement principles, they are **social values** that have meaning and force outside of measurement wherever evaluative judgments and decisions are made. (p.13)

With regard to special claims of “authenticity”, Messick notes,

The portrayal of performance assessments as **authentic** and **direct** has all the earmarks of a validity claim with little or no evidential grounding. That is, if authenticity is important to consider when evaluating the consequences of assessment for student achievement, it constitutes a tacit validity standard, as does the closely related concept of directness of assessment. We need to address what the labels authentic and direct might mean in validity terms. We also need to determine what kinds of evidence might legitimize both their use as validity standards and their nefarious implication that other forms of assessment are not only indirect, but inauthentic. (p.14)

As previously noted, Wiggins presents no validity data, evidential or consequential, in any of his writings on “authentic” assessment.

Assessment and Educational Philosophy

A theme that runs through Wiggins' conception of “authentic” testing is an emphasis upon performances that are designed to assess “higher order” outcomes. Among the design features he lists (1989a,p.45) as characteristic of “authentic” tests are: contextualized, complex intellectual challenges; involves student's own research; emphasizes **depth** more than breadth; and involves somewhat “ill-structured” tasks or problems. He specifically disavows any interest in “atomized” tasks, corresponding to isolated “outcomes”, “mere recall”, and “plug-in skills”. The reason for this is made clear in his discussion of validity in his 1993 book.

At bottom is a philosophical problem of major proportions about the purpose of schooling: Is schooling meant to yield common knowledge? If so, then it makes perfect sense to think of tests as properly focusing on what students hold in common. But what if education is a personal, idiosyncratic affair, where the meaning and personal effectiveness that I derive from coursework is more important than what knowledge we all end up holding in common? In that case, a standardized, indirect test- -of any kind- -would make no sense: What could we possibly mean by a standardized test of meaning of educative experience? What, then, of the validity of aptitude tests if I cannot state with precision: aptitude **for what future role or aspiration?** (p. 247)

Wiggins (1993) gives examples of a variety of “roles and situations through which students can perform with knowledge”. The roles include such diverse occupations and settings as follow (p.223-24):

- | | |
|--------------------|-----------------------------|
| museum curator | engineering designer |
| U.N representative | ad agency director |
| tour organizer | psychologist/sociologist |
| bank manager | newspaper editor and writer |
| historian | product designer |

job applicant	teacher
expert witness	speaker-listener
reviewer	commercial designer

He argues that such roles could serve as “template” for better test design because, “These roles and situational challenges are common to professional life.” Wiggins claims it is a “logical fallacy” to argue that students must be given “drills” and “tests concerning their mastery of the drills” prior to requiring performance in professional roles and situations. “Drill testing” is a means to an end and “it is certainly not to be confused with the important performance itself”.

Wiggins conveniently ignores the possibility that most ‘roles and situational challenges common to professional life’ involve an extensive knowledge base. Individuals who lack the knowledge base have little or no chance of performing successfully in the ‘real-life’ roles which he describes. For example, an historian typically specializes in a particular time period and geographical region for her/his research, e.g. Colonial America during the period 1650-1770. In order to make a useful contribution to the literature on this topic, the historian must first become familiar with a vast amount of work previously published by other historians working on this and related topics.

The Importance of Knowledge

Knowledge and the Taxonomy

I believe that phrases like “drills” and “drill testing” are employed with the intent of denigrating the role of **knowledge** in the various “roles” and “situational challenges” to which Wiggins refers. In fact, knowledge is a basic building block upon which most “higher order” educational outcomes rest. This is explicitly acknowledged in the most widely cited framework for classifying educational outcomes, the Bloom taxonomy, which defines “knowledge” as the most elementary of all the cognitive outcomes.

It is interesting to note that Wiggins claims that the authors of the taxonomy did not intend to imply that educational outcomes necessarily occur in a “fixed chronological sequence” beginning with knowledge. This claim is made despite the fact that the taxonomy is clearly presented by its authors as an **hierarchical** model with outcomes arranged on a simple- to-complex scale of cognitive processes.

It is also important to recall that “knowledge” as employed by the authors of the taxonomy extends well beyond conventional notions of “facts”. Subcategories in the taxonomy include the following:

Knowledge of terms.
Knowledge of facts.
Knowledge of concepts.
Knowledge of principles.
Knowledge of procedure.
Knowledge of theories.

This list of diverse outcomes should make it clear that the “knowledge” intended by the authors of the taxonomy is not necessarily limited to the type of “knowledge” achieved strictly through “drills” and “drill testing”

Knowledge and Expertise

The fundamental role played by knowledge in “real-life” roles is well documented in the extensive body of work on the nature of expertise. Chi, Glaser, and Farr (1988) have edited a series of papers which summarize much of the work in this field. In their overview of this work Glaser and Chi list several “key characteristics” of the performance of experts. The first characteristic they cite is that expertise generally is restricted to specific domains of performance and does not transfer from one domain to another. They state,

The obvious reason for the excellence of experts is that they have a good deal of domain knowledge. This is easily demonstrated; for example, in medical diagnosis, expert physicians have more differentiations of common diseases into disease variants (Johnson et al., 1981). Likewise, in examining taxi drivers' knowledge of routes, Chase (1983) found that expert drivers can generate a far greater number of secondary routes (i.e. lesser known streets) than novice drivers. (1988, p. xvii)

Knowledge and Literacy

Wiggin's emphasis upon roles “common to professional life” reflects a very narrow view of educational outcomes. Surely, there are many other outcomes of education which are also important, e.g. actively participating in the economic, political, and cultural life of one's community; acting as a responsible member of a family and various other social organizations; making contributions to the improvement of the life of one's community through involvement in civic and volunteer organizations, etc. In other words, it is also important that students develop into adults who take their roles as citizens seriously and who make positive contributions to the improvement of society.

Hirsch (1987) argues that a fundamental goal of education should be to produce students who are “culturally literate”. His bestselling book, **Cultural Literacy-What Every American Needs to Know**, was written in response to several studies which indicated a serious decline in the level of knowledge possessed by American students during the 1970's and early 1980's. Hirsch describes cultural literacy as the “network of information that all competent readers possess”.

It is the background information stored in their minds, that enables them to take up a newspaper and read it with an adequate level of comprehension, getting the point, grasping the implications, relating what they read to the unstated context which alone gives meaning to what they read. (1987,p.2)

Hirsch is critical of attempts to dismiss knowledge in favor of more lofty educational goals. He argues that the denigration of “mere facts” by advocates of instruction in “higher order” skills creates a false dichotomy.

The polarization of educationists into facts-people versus skills-people has no basis in reason. Facts and skills are inseparable. There is no insurmountable reason why those who advocate the teaching of higher order skills and those who advocate the teaching of common traditional content should not join forces. No philosophical or practical barrier prevents them from doing so, and all who consider mature literacy to be a paramount aim of education will wish them to do so. (1987, p.133)

Proponents of educational reform who stress “critical thinking” and similar “higher order” thinking skills should consider Hirsch's advice seriously when planning classroom instruction and assessment.

The Price of Ignorance: A Recent Warning

(This section is based largely upon an article by Richard Morin of the **Washington Post** which appeared in the **Minneapolis Star Tribune** on February 12, 1996.)

There are many studies which document the lack of knowledge that characterizes American students on topics ranging from math and science to history and geography. A recent survey conducted by the **Washington Post**, the Kaiser Family Foundation, and Harvard University reveals a similar lack of knowledge in the area of U.S. government and politics. The study authors interviewed 1,514 randomly selected adults (not students) in November and December, 1995. Study participants were asked 18 general knowledge questions about the government and leaders in government. An additional 21 questions concerning political knowledge were asked in four other national surveys conducted by the **Post**. Some selected findings are as follow:

two-thirds of those interviewed could not name their representative in the U.S House of Representatives

40% did not know the name of the Vice President

two-thirds could not name the majority leader of the U.S. Senate

46% did not know the name of the speaker of the House of Representatives

75% were not aware that U.S. senators are elected for six- year terms

40% did not know that Republicans control both chambers of Congress
nearly 60% incorrectly believed that the government spends more on foreign aid than Medicare

The survey findings further suggests some of the more important consequences of lack of knowledge. Those who are less informed are much more likely to tune out politics and turn off voting. Also, those who are less informed are more likely to believe that the country is in decline. For example, less informed Americans are far more likely to believe that that annual budget deficit and the number of federal employees had increased- - not decreased--in recent years.

The gap in information also affects how politics is practiced according to Samuel Popkin, a political scientist at U.C., San Diego. He states that candidates for office run two distinct campaigns; one for informed voters which stress their stands on issues and policies and one in which strategists and consultants aim attacks at the character of the opponent to win support from less-informed voters. Popkin claims that , for the less- informed voter, all politics is reduced to character and caricature: politicians are separated into "heroes and villains"and major policy debates become clashes between good and evil. The existence of a substantial block of poorly informed voters is one reason why so-called 'negative campaigning' seems to be so effective

A Case History of an "Authentic" Test

The notion that complex measures of achievement be used to assess performance in more a "lifelike" setting than is represented by conventional standardized tests is not new. In fact, instruments designed to be more "authentic" than paper-and - pencil exams have been employed in particular occupational and educational setting for many years .¹ The in-basket test is the product of over five decades of research and application in business and managerial settings. It is a simulation exercise using materials typically found in a hypothetical managerial position. The materials in the in-basket consist of letters, memos, records, and other items which require a response from the examinee who role-plays as the occupant of the hypothetical position. The in-basket methodology is frequently used as an assessment tool for predicting job performance.

¹ Swanson, Norman, and Linn (1995) recently reviewed the long history of such measures in the health professions. Their article describes four specific performance-based methods commonly employed in training health-care professionals and lists "some hard lessons learned from many studies and frequent missteps". Advocates of the use of complex performance measures in large scale assessment programs would be well advised to study the eight "lessons" listed in this article.

Schippmann, Prien, and Katz (1990) did a comprehensive review of the literature on the psychometric properties of in - basket tests. They examined all available studies that reported reliability and/or validity data using in- basket tests. Their findings are as follow:

Interrater reliability-- "it appears that scorers and raters are responding fairly consistently to participant response data. The range of interrater values, however, suggests that something else is operating which creates deviant score/rating patterns. This 'something else' may well be a function of scorer/rater training."

Alternate-form reliability--"The difference between obtained alternate-form reliabilities and interrater reliabilities is dramatic. In the absence of additional research data, the conservative conclusion is that the consistency of individual performance is, at best, marginal."

Split-half reliability--"split-half reliabilities of in- basket performance measures have been somewhat disappointing. Though it would appear that greater rigor in the development of test content and more systematic and objective scoring procedures might yield more encouraging reliabilities, additional research is needed in order to draw firm conclusions."

Content validity--"all of the reviewed studies which suggest that their procedures are content valid fall seriously short of the mark in our view."--"There simply are no published or widely distributed reports which describe how to develop an in- basket that is well grounded with regard to content validity."

Criterion-related validity--"studies of criterion-related validity did reveal a large number of significant correlations between in-basket measures and various criteria. Thus the evidence of criterion-related validity for certain in-baskets is sufficient to support the development and use of the procedure for various decision making purposes."--"However, it is frequently difficult to determine the in-basket's contribution to the prediction of performance apart from other assessment tools."

Construct validity--"evidence of construct validity was suggested in those cases where the in-basket was designed to reflect some theoretical construct. The conclusion warranted here is that the evidence from these studies is encouraging, though not convincing with reference to either hypothetical constructs or job performance criterion constructs."

Schippmann, et al. (1990) make some observations in their discussion of validity that have special relevance for advocates of "authentic" tests.

the issue of the content validity of in-basket exercises appears to be clouded by confusion between face validity and content validity. Simply eyeballing test content with reference to the job analysis results and proclaiming the procedures to be content valid is unacceptable. The appearance of the relevance may contribute to face validity, though it is not a sufficient basis for making inferences about job performance . (p.851)

Conclusions

Wiggin's promotion of "authentic" assessment, however well intentioned, is flawed in several important respects. First, the term "authentic" is misleading and confusing. The term inappropriately implies that some assessment approaches are superior to others because they measure outcomes that are more "genuine" or "real". This claim is based largely upon an appeal to "face" validity, a concept which has been abandoned by modern psychometric theorists. Second, Wiggins rejects the role of knowledge in the assessment of educational outcomes. This ignores a substantial body of theory and ample empirical evidence which supports the central role of knowledge in countless domains of performance. Third, the presumption that complex performance-based measures of assessment can be designed to achieve the psychometric quality necessary for use in large-scale assessment programs is, at best, questionable in light of previous efforts to employ such measures in other professions.

References

- Darling-Hammond, L., Anness, J., and Falk, B. (1995) *Authentic Assessment in Action: Studies of Schools and Students at Work*. New York: Teacher College Press.
- Chi, M., Glaser, R., and Farr, M. (Eds.) (1988). *The Nature of Expertise*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Fitzpatrick, R. and Morrison, E. (1971). Performance and Product Evaluation. In R.L. Thorndike (Ed.) *Educational Measurement*, 2nd Ed. New York: American Council on Education/Macmillan.
- Hirsch, E.D. (1987). *Cultural Literacy: What Every American Needs To Know*. New York: Houghton-Mifflin Co.
- Lesh, R. and Lamon, S. (Eds.) (1992) *Assessment of Authentic Performance in School Mathematics*. Washington: AAAS Press.
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessment. *Educational Researcher*, 23 (2), 13-23.
- Morin, R. (1996). Of politics, government: Citizen ignorance rampant, costly. *Minneapolis Star Tribune*, pg. A5, Feb. 12.
- Schippman, J., Prien, E., and Katz, J. (1990). Reliability and validity of in-basket performance measures. *Personnel Psychology*, 43 (4), 837-859.

Swanson, D., Norman, G., and Linn, R. (1995) Performance-based assessment: Lesson from the health professions. *Educational Researcher*, 24 (5), 5-11,35.

Wiggins, G. (1998a). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan*, 20, 703-713.

Wiggins, G. (1989). Teaching to the (authentic) test. *Educational Leadership*, 46. 41-47.

Wiggins, G. (1993). *Assessing Student Performance*. San Francisco: Jossey- Bass Publishers.

The Validity and Reliability of Information from a Student Annotation Form for Work Samples in Portfolios

Susan M. Brookhart*

ABSTRACT

This paper presents evidence for the validity and reliability of information from a Student Annotation Form used to collect student reflections on work samples in portfolios for use in a districtwide evaluation. The forms were designed to be general enough to apply to many grades and subjects, so that they could be used across classes in the evaluation. The annotation forms were simple enough for students at all levels to complete. The forms asked students to indicate how difficult they found the work sample to be, whether they would like to do more of the work, and why. The difficulty and do-more questions were multiple choice; the "why" question was open-ended. For each work sample that illustrated a curriculum objective, selected for inclusion in an evaluation portfolio, students completed an annotation form and attached it to the work sample. Study One was performed with data collected in 1994, from a total of 367 students, grades 1 through 10, in whose portfolios were included a total of 1678 annotated work samples. The validity and reliability of the difficulty and do-more judgments were confirmed. A constructed measure of academic self-efficacy looked promising at first but failed a validity check. Study Two was performed on data collected in 1995, from a total of 313 students, grades 1 through 10, in whose portfolios were included a total of 2862 annotated work samples. The quantitative reliability and validity results were replicated for the difficulty and do-more questions.

* Duquesne University

A version of this paper was presented at the 1996 Annual Meeting of the American Educational Research Association. This work was supported in part by a Duquesne University Presidential Summer Writing Scholarship, and in part by the Fox Chapel School District. The analysis and conclusions expressed here are the responsibility of the author. Thanks go to Laurel J. Swenson for her assistance with this study.

Various kinds of portfolios, constructed according to all sorts of different rules, are being used to assess student achievement, progress, attitudes, and understandings for many different purposes, including regular classroom assessment, evaluation of school programs, and evaluation of district-level achievement. One common element in all portfolios, an element that distinguishes them from traditional measures of achievement and, some say, is the key to their usefulness, is the opportunity they offer for students' reflections on their work.

Many forms, procedures and strategies have been suggested for student reflection on portfolio work. These include checklists of work done and reflection questions to answer (Farnan & Kelly, 1991), letter-writing or summarizing assignments (Ferguson & Maples, 1992), and various designs of self-evaluation forms, for annotating individual work samples or for reflecting on the whole portfolio (Azar et al., 1993; Goodman, Bird, & Goodman 1992). There is some evidence that teachers' use of student reflection increases as they work with portfolios over time (Roe & Vukelich, 1995). Reliability and validity studies have typically been performed on the ratings of achievement or performance in portfolios (e.g., Klein, McCaffrey, Stecher, & Koretz, 1995), however, not on student reflections. The results of these studies of the technical qualities of portfolio achievement measures have been mixed (Calfee & Perfumo, 1993; Herman & Winters, 1994). No studies were found on the technical qualities of portfolio measures of student reflections or dispositions toward their work.

This paper presents evidence for the validity and reliability of information from a student annotation form used to collect student reflections on work samples in portfolios for use in a districtwide evaluation. The forms were designed to be general enough to apply to many grades and subjects, so that they could be used across classes in an evaluation project. Part of a larger evaluation project for the Fox Chapel (PA) Area School District required the design of achievement and student progress measures that tapped some outcomes of instruction not measured by standardized tests or by teacher judgment, since these data are readily available in other measures. Specifically, the measures were to tap student judgments about their academic work and themselves as academic workers. These kinds of judgments contribute to academic self-efficacy (Zimmerman & Martinez-Pons, 1992), and they are an important and poorly-measured aspect of what the general and political public at present calls "21st century skills." Beyond problem-solving ability *per se*, there are orientations that will genuinely establish today's students as tomorrow's problem-solvers. Not everyone who can solve a problem, does. The efficacy measures in this evaluation were to tap feelings of efficacy at doing the particular work

sampled in the portfolio, not the more general academic self-efficacy for which standardized measures are available

One of the measures designed was a Student Annotation Form, to be completed and attached to work samples in student portfolios. For each work sample that illustrated a curriculum objective, selected for inclusion in an evaluation portfolio, students completed an annotation form and attached it to the work sample. The design of the Student Annotation Form was influenced by previous experience in a pilot study. The pilot study had used a more open-ended annotation form that asked, "What did you learn?" The data gathered with this form had been more likely to include the title of a page of work, e.g., "Two-digit multiplication," than anything reflective. When a student copied the assignment title, it was not certain that the student knew what it meant. Thus the Student Annotation Form (see Figure 1) asked students to make two multiple-choice judgments, about the difficulty of the work sample and whether or not they would like to do more of that kind of work, and to write briefly "Why" they would or would not like to do more.

Figure 1 Student Annotation Form (Student copies were half-sheets to be stapled to work sample)

Circle your choices and then tell why, Staple this paper to your work sample	
I found this work	a. easy b. about the same as other work. c. hard
I would like to do more of this kind of work	a. yes b. no
Why?	

The expectation was that in responding to the "Why" question, students would tip their hands about what they thought the assignment meant, at least for themselves. The choices for student judgment of difficulty were based in common classroom talk. Many students say a task was "easy" when they mean to boast "I could do it, and I did it successfully." In the experience of the evaluation team who developed the Student Annotation Form, a student judgment that work was "easy" did not necessarily mean that it was beneath their capabilities, at the level of unchallenging busywork, but it was a judgment they made naturally. The opposite of "easy," for students, is "hard"

The use of multiple-choice questions to stimulate reflection is not recommended as a strategy for portfolios intended to be used primarily for instruction. The intended use of information here, however, was as an indicator of student perceptions of work done under a program called Continuous Progress Instruction, based on Wang's (1992) Adaptive Learning Environments Model. Most of the classes participating in the initial implementation phase were mathematics classes, but some were English classes. Student age and developmental level varied widely, so the multiple choice items were viewed as basic indicators for the evaluation that all students could do and that would lead some students to deeper reflection in the open-ended responses that followed. This study is concerned with validating the use of the student annotation form for this cross-classroom evaluation.

Student effort and achievement has been shown to be related to both perceptions of task characteristics and students' belief in their abilities to handle those task characteristics (Salomon, 1984). Therefore, the evaluators experimented with a measure of self-efficacy to do the particular kind of work demonstrated in the portfolio. The judgment of difficulty (DIFF=1 if the work was judged easy, 2 if about the same as other work, and 3 if hard) was multiplied by the willingness to do more of that particular kind of work (DOMORE=1 if yes, 0 if no). Thus the self-efficacy indicator (SE) ranged from 0 (would not like to do more) through 1 (would like to do more of this easy kind of work), 2 (would like to do more of this average kind of work), to 3 (would like to do more of this hard work). It was reasoned that work a student was unwilling to do would not contribute to a student's image of himself or herself as a worker in that area and that as students were willing to tackle work they judged progressively more difficult, they would have progressively stronger images of themselves as workers in that area. In addition to results for the validity of information from the two items on the student annotation form, this study presents results of a validity check on the constructed measure of self-efficacy. Briefly, the items on the form itself were found valid for their intended districtwide evaluation purpose, but the experimental constructed measure of work sample-dependent self-efficacy was not found to be valid.

Research Question

The annotation forms designed for the Fox Chapel evaluation offer a very simple and widely applicable tool for student reflection. But before this tool is widely applied, it is very important to establish that the information elicited by these forms is valid and reliable, that is, that the responses are accurate and dependable representations of what the students think. The research question under investigation was as follows: What evidence is there for the validity and reliability of student annotations of work samples in portfolios?

Specific to validity, do student's ratings of the difficulty of the work match their perceived ability to learn the material? When a student reports wanting to do more of the work, what reasons are given? Are these reasons conducive to learning? Does the constructed self-efficacy variable (formed by multiplying the difficulty and do more responses) match students's perceived ability to learn the material?

Specific to reliability, what is the internal consistency of the set of annotations in each student's portfolio? What is the generalizability of responses across time (as the students use the same annotation form over and over)? Internal consistency and generalizability were important to investigate because in the evaluation studies, the annotations were aggregated across portfolios to arrive at student overall ratings of the difficulty of their work and their willingness to do more. These uses required that the annotations may be aggregated reliably. For using information from an individual annotation and its work sample in classroom instruction, the validity questions about the relationship between student reasons for ratings and learning, listed above, would be of primary importance.

Study One includes results of a complete investigation of these research questions using data collected in 1994. Reliability and validity of information from the DIFF, DOMORE, and SE variables were studied with both quantitative and qualitative methods. Qualitative analysis was required for the open-ended responses to the "Why?" question. Study Two includes results of a replication of the quantitative analyses of the DIFF and DOMORE variables, using data collected in 1995. Analyses of the SE variable were eliminated because of the results of Study One.

Study One

Method

Data. Data were collected in 1994. Within each class in the evaluation project, portfolios from 10 students were selected according to a stratified (gifted, nonlabeled, and educational support) random sample. There were portfolio data from a total of 367 students, grades 1 through 6 and grade 10, in whose portfolios were included a total of 1678 annotated work samples. There were students from every school in the district: 4 elementary schools, one middle school (the 6th graders) and the high school. Most of the portfolios were from mathematics classes (82%); the rest were from English/Language Arts classes (18%). Fifty-four percent of the students were male; 46% were female.

Student status in this sample was as follows: gifted, 23%; nonlabeled, 58%; educational support, 19%. All classes in the sample were part of the initial districtwide implementation of Continuous Progress Instruction.

Data included student responses to each Student Annotation Form (number of annotated work samples per student portfolio ranged from 0 to 13) in the portfolio and teacher ratings from a cover sheet, preprinted with appropriate curriculum objectives for the class, that the teachers inserted in each student's portfolio. Teacher ratings of achievement used a 1-4 rubric: The student demonstrates (1) no evidence of mastery of the objective, (2) partial/incomplete mastery, (3) acceptable mastery, or (4) exceptional mastery. Teachers rated efficiency of use of study time for each curriculum objective demonstrated in the portfolio (1=efficient, 0=not efficient). Background information (grade, subject, school, student status, and gender) about students was included in the data set. Mean DIFF, DOMORE, and SE ratings were calculated for each student's portfolio, based on the total number of annotations in that portfolio. Missing data resulted in varying sample sizes for each analysis; actual n for each analysis is reported in the tables.

Analyses. Two reliability analyses were done. (1) Cronbach's alpha was calculated for each rating, treating annotations within one student's portfolio as items on DIFF, DOMORE, and SE scales, respectively. Student was the unit of analysis. Internal consistency reliability is required if meaningful composite DIFF and DOMORE ratings for each student are to be calculated. Overall mean DIFF and DOMORE ratings were used in the evaluation, to answer evaluation questions about students' reported dispositions toward their work.

(2) Generalizability studies were conducted for each of the three scales, with the design Person by Time. The Time facet represented the ordinal position of the annotation in the student's portfolio: the first time the student used the form, the second time, and so on. This design examined both generalizability across time and the size of the Person-by-Time interaction. Generalizability studies were conducted to answer two reliability questions. First, how many annotations per portfolio are required for a reliable estimate of students' overall dispositions toward their work? Second, does using the same form over and over result in patterns of responses that might indicate students tire of reporting DIFF and DOMORE? If so, questions are raised about the validity of continued use.

Additional quantitative and qualitative validity analyses were conducted. Correlations of DIFF, DOMORE, and SE with overall teacher ratings of achievement, effectiveness of use of study time, and percent of curriculum objectives attained were examined.

Mean ratings of achievement and effectiveness of time use for each student were calculated over all curriculum objectives that the student attempted; number and level of objectives attempted differed for each student. Percent of objectives attained was calculated as the number of objectives on which a student received as satisfactory or better teacher rating, divided by the total number of objectives marked on the student's portfolio cover sheet. The student was the unit of analysis. These correlations provided a look at the external structure of information from the Student Annotation Forms. The external measures were teacher ratings of achievement and time use on the same curriculum objectives for which work samples were selected and annotated for the portfolios.

Two qualitative analyses of the open-ended responses addressed these research questions: Do students' ratings of the difficulty of the work match perceived ability to learn the material? When a student reports wanting to do more of the work, what reasons are given? Are these reasons conducive to learning? What reasons for wanting to do more of the work are given at each level of self-efficacy, and are these reasons consonant with the interpretation of academic self-efficacy the experimental variable is thought to measure? Individual annotations were the unit of analysis. First, student responses to the open-ended question "Why?" were coded according to how the student viewed further learning; (a) response contained reasons or opinions that would facilitate further learning for the student, (b) response contained reasons or opinions that would hinder further learning, (c) neutral or ambiguous, and (d) response contained reasons or opinions that would both facilitate and hinder further learning for the student. To check the reliability of coding, two independent coders categorized a random sample of 100 of the written comments into these four categories; their agreement rate was 87%. Table 1.1 illustrates these codes with some example student responses from each category.

The second analysis examined key words in the text of students' written annotations. The words "easy," "fun," and "hard" occurred often. These words are also commonly used in students' talk about schoolwork. Word searches on each of these words were done by computer, then augmented by visual inspection of the text. Inspection allowed for adding misspellings and invented spellings that the computer did not find, for example, "ese" for "easy" or "hrd" for "hard," and to add variations that were not found in the computer search, for example, "easier." Inspection also allowed the separation of "too easy," a negative response, from "easy," which students usually use as a positive description of schoolwork. If a word was repeated within one annotation, it was counted only once for example, the comment because it is easy and I like easy work " was counted as one expression of "easy." Responses were organized by DIFF, DOMORE, and SE values, then a table was created tallying word use at each rating level.

Table 1.1 Examples of Student Written Comments about Why They Would or Would not Like to Do More of the Type of Work Represented by a Work Sample

Sample Comment	Learning Code
Because I got a better understanding of the book and I saw what other people thought of the book.	1, facilitates learning
Because using division patternes are easy and hard, and if I do more I could probably get a A on my test.	1, facilitates learning
Because I don't like word problems	2, hinders learning
I thank it is the wors papr in the world and I do not want to do it agane !!!!!	2, hinders learning
A different y way to learn	3, neutral or ambiguous
I guess that this problem is ok, I still think it should explain more. I just don't like the extra enrichment. It takes to long. But it lets us go at our own rate.	4, both facilitates & hinders

Results and Discussion

Reliability. In general, reliability was acceptable for the districtwide evaluation purpose. Alpha reliability for seven annotation forms per portfolio was .62 for difficulty ratings, .71 for the domore ratings, and .69 for academic self-efficacy. Table 1.2 presents alpha values for 6 through 9 annotations per portfolio. Very few of the portfolios included more than 9 annotated work samples. Internal consistency of annotation of work samples within a portfolio would clearly depend in part on the similarity of the work samples to one another; it was not expected that the annotations would be as internally consistent as the items on a test or survey.

Table 1.2 Reliability (Internal Consistency) of Scales, for Six through Nine Annotations per Portfolio

Scale	k	α	n	α'
DIFF	6	.59	158	.71
	7	.62	121	.70
	8	.60	53	.65
	9	.64	20	.66
DOMORE	6	.66	157	.76
	7	.71	118	.78
	8	.73	53	.77
	9	.80	19	.82
SE	6	.65	153	.76
	7	.69	116	.76
	8	.74	51	.78
	9	.84	19	.85

α' = Reliability estimate for k = 10, using the Spearman-Brown formula

If reliability is considered to be the usefulness of the annotations across time, generalizability is at issue. As students use the same annotation form over and over again, for successive work samples included in their portfolios, do their judgments waver? For the DOMORE and SE variables, in fact, the variance due to the Time factor was zero (see Table 1.3), leading to identical values for both absolute and relative generalizability in the Person by Time design studied. The Person-by-Time interaction for each was large (about 80% of total variance). This means that different students did change the difficulty and do-more ratings in somewhat different ways over time; in part, this reflects the fact that work samples themselves differed. The large Person-by-Time interaction also showed that the student annotation forms collected data that behaved in a similar fashion to most performance assessment data, which typically has a large person-by-task variance component (Brennan, Gao, & Colton, 1995). So while one of the challenges of instrument development will be to reduce this term, at present its size suggests that students approach the student annotation forms in a way more similar to how they would approach performance assessments than paper-and-pencil surveys. All three scales generalized at an acceptable level across uses (see Table 1.4). Generalizability coefficients for these three ratings were similar.

Table 1.3 Variance Component Estimates for Scales, for Person x Time Design

Source of Variability	DIFF		DOMORE		SE	
	Estimated Variance Component	Percent Total Variability	Estimated Variance Component	Percent Total Variability	Estimated Variance Component	Percent Total Variability
Person	.0913	19%	.0527	26%	.1640	24%
Time	.0036	1%	.0000	0%	.0000	0%
P x T	.3923	81%	.1495	74%	.5242	76%

DIFF & DOMORE variance estimates based on 118 persons, 7 times

SE variance estimates based on 116 persons, 7 times

Table 1.4 Generalizability of Scales, for Six through Ten Annotations per Portfolio

Scale	k	$\hat{\rho}^2$	$\hat{\phi}$
DIFF	6	.58	.58
	7	.62	.62
	8	.65	.65
	9	.68	.67
	10	.70	.70
DOMORE	6	.68	.68
	7	.71	.71
	8	.74	.74
	9	.76	.76
	10	.78	.78
SE	6	.65	.65
	7	.69	.69
	8	.71	.71
	9	.74	.74
	10	.76	.76

DIFF & DOMORE generalizability estimates based on 118 persons

SE generalizability estimates based on 116 persons

Validity. As the rationale for these annotation forms suggests, it was expected that these forms would tap self-efficacy about particular schoolwork by indicating student judgments about the difficulty of the work and their willingness to pursue it. Table 1.5 shows there were weak but significant correlations between students' mean DIFF ratings and teacher judgment of how efficiently overall students used their time in independent work. ($r=-.17, p=.005$) and between students' mean DOMORE ratings and teacher judgment of how efficiently students used their time ($r=.12, p=.05$). Teachers judged more efficient time use for students who rated their work samples less difficult overall. Teachers judged more efficient time use for students who indicated more interest in doing their work. This weak convergent evidence suggested that students' in-class behavior was somewhat consistent with their annotations.

Table 1.5 Correlations between Scales and Selected External Measures

	PCTACH	TIME	EVAL
SE	-.02	-.01	.04
DIFF	-.09	-.17**	-.10
DOMORE	-.02	.12*	.03

* $p=.05$, ** $p=.005$, n ranged from 269 to 286 students

PCTACH = percent of learning objectives attempted within the portfolio that were achieved

TIME= efficiency of time use, 0=inefficient, 1=efficient, averaged over all learning objectives represented in the portfolio

EVAL= teacher evaluation of achievement, on a scale of 1 to 4, averaged over all learning objectives represented in the portfolio

None of the three scales (DIFF, DOMORE, or SE) was significantly correlated with the percent of attempted learning objectives, documented in the portfolio, that were achieved (each portfolio was supposed to include one work sample per learning objective), as measured by teacher judgment, or with teacher judgment of the level of achievement indicated in the portfolio. This divergent evidence suggested that the student annotation forms were not measuring the same thing as teacher judgment of achievement

Cross- tabulations between DIFF, DOMORE, and SE scales (Tables 1.6, 1.7, and 1.8, respectively) and the four LEARNING categories showed expected relationships. For example , for annotations on which the students responded yes, they would like to do more of the kind of work in the sample, 91% of the written responses were about the work facilitating further learning; conversely, for annotations on which the students responded no, they would not like to do more of that kind of work, 73% of the written responses gave evidence of hindering further learning (Table 1.7)

Table 1.6 Crosstabulation of Students' Judgments of Difficulty with Their Judgments about the Value of the Sampled Work for Their Learning

LEARNING Frequency (Col Pct)	DIFF			Total
	EASY	SAME AS OTHER WORK	HARD	
FACILITATES	674 (77%)	326 (64%)	68 (33%)	1068 (67%)
HINDERS	134 (15%)	125 (25%)	118 (57%)	377 (24%)
NEUTRAL/ AMBIGUOUS	36 (4%)	27 (5%)	13 (6%)	76 (5%)
BOTH	36 (4%)	29 (6%)	9 (4%)	74 (5%)
Total	880	507	208	1595

Unit of analysis = annotation of one work sample

LEARNING was code from student responses to open-ended question, "Why?"

Table 1.7 Crosstabulation of Students' Willingness to Do More with Their Judgments about the Value of the Sampled Work for Their Learning

LEARNING Frequency (Col Pct)	DO MORE		Total
	NO	YES	
FACILITATES	67 (14%)	1003 (91%)	1070 (67%)
HINDERS	356 (73%)	26 (2%)	382 (24%)
NEUTRAL/ AMBIGUOUS	26 (5%)	48 (4%)	74 (5%)
BOTH	40 (8%)	28 (3%)	68 (4%)
Total	489	1105	1594

Unit of analysis = annotation of one work sample

LEARNING was coded from student responses to open-ended question, "Why?"

Table 1.8 Crosstabulation of Student's Self-Efficacy about the Work with Their Judgments about the Value of the Sampled Work for Their Learning

LEARNING Frequency Col Pct	SELF-EFFICACY				Total
	0	1	2	3	
FACILITATES	67 (14%)	630 (93%)	306 (89%)	61 (78%)	1064 (67%)
HINDERS	356 (73%)	9 (1%)	8 (2%)	8 (10%)	381 (24%)
NEUTRAL/ AMBIGUOUS	26 (5%)	25 (4%)	15 (4%)	7 (9%)	73 (5%)
BOTH	40 (8%)	12 (2%)	14 (4%)	2 (3%)	68 (4%)
Total	489	676	343	78	1586

Unit of analysis = annotation of one work sample

SE=DIFF X DOMORE

LEARNING was coded from student responses to open-ended question, "Why?"

Correlations of mean DIFF, DOMORE, and SE for each student with the mean LEARNING indication in a student's written comments (transformed into 1=facilitates, 0=hinders, disregarding ambiguous responses) indicated that self-efficacy for learning and wanting to do more of a particular kind of work were positively related to perceptions of usefulness for learning, while perceived difficulty level was negatively related to perceptions of usefulness for learning. This is consistent with cognitive theory that says students need a moderate level of challenge. These relationships held, in the same strength and direction, for gifted, nonlabeled, and educational support students, with the exception of the relationship of DIFF to LEARNING for gifted students (see Table 1.9). These relationships also held for the most part across grade levels; the exceptions were grades 2 and 5 (see Table 1.10).

Usage of key words from the students' own writing are organized in Table 1.11 according to the DIFF, DOMORE, and SE values for the annotation on which the comment appeared. The frequencies and percents in the table may be considered a conservative estimate of usage. To keep the search from requiring much inference on the part of the reader, only the key words were counted. Invented spellings were allowed, but possible conceptual connections were not. For example, one student wrote "I already no [know] it" and might well have considered the work easy or too easy, but the comment was not tallied as an instance of "easy." The only conceptual inference permitted in the search was the removal of negatives as an instance of usage. For example, one student wrote, "I'd rather do something fun," and this was not counted as an instance of fun, since the student clearly did not think this work was fun.

Table 1.9 Correlation between DIFF, DOMORE, and SE Scales and Students' Judgments of the Value of the Work to Their Learning, Overall and by Student Status

n	All Students 308	Gifted 68	Status	
			Nonlabeled 185	Educational Support 55
DIFF	-.22	.02	-.28	-.36
DOMORE	.80	.81	.76	.93
SE	.63	.73	.57	.74

Unit of analysis = student

Learning judgment is the average, of all annotations in a student's portfolio, of written responses to "Why" s/he would do more of the kind of work in the sample, coded 1 = facilitates learning, 0 = hinders learning (ambiguous responses were not used in this analysis).

Table 1.10 Correlation between DIFF, DOMORE, and SE Scales and Students' Judgments of the Value of the Work to Their Learning, Overall and by Grade

	All Students	Grade						
		1	2	3	4	5	6	10
n	308	49	52	55	27	59	21	45
DIFF	-.22	-.34	.21	-.55	-.12	.01	-.28	-.32
DOMORE	.80	.80	.93	.95	.63	.64	.75	.85
SE	.63	.59	.80	.61	.48	.56	.67	.73

Unit of analysis = student

Learning judgment is the average, of all annotations in a student's portfolio, of written responses to "Why" s/he would do more of the kind of work in the sample, coded 1 = facilitates learning, 0 = hinders learning (ambiguous responses were not used in this analysis).

Table 1.11 Content Analysis of Frequently Used Words from "Why?" Responses on Student Annotation Forms

Ratings	Total	Word(s)			
		Too easy	Easy	Fun	Hard
SE = 0	482	58 (12%)	56 (12%)	4 (1%)	93 (19%)
DIFF = 1, DOMORE = 0	199	53 (27%)	34 (17%)	3 (2%)	10 (5%)
DIFF = 2, DOMORE = 0	157	4 (3%)	17 (11%)	1 (0%)	19 (12%)
DIFF = 3, DOMORE = 0	126	1 (0%)	5 (4%)	0 (0%)	64 (51%)
SE = 1	677	1 (0%)	238 (35%)	194 (29%)	8 (1%)
DIFF = 1, DOMORE = 1					
SE = 2	341	0 (0%)	41 (12%)	93 (27%)	15 (4%)
DIFF = 2, DOMORE = 1					
SE = 3	77	0 (0%)	3 (4%)	12 (16%)	16 (21%)
DIFF = 3, DOMORE = 1					

Values in the table are frequencies of occurrence of each word, and percent of total usage for each rating level of the work

The information in Table 1.11 suggested that usage of "easy" and "hard" varied as expected with DIFF ratings. This supported the validity of these ratings. Observations about key word usage by SE indicated that for SE=0 (and thus for DOMORE=0), "fun" was conspicuously absent. For SE=1 and SE=2, "easy" and "fun" work was more common. For SE=3, "hard" work was more common.

Visual inspection of the comments suggested that the tone of the instances of "hard" for SE=0 was very different from the instances for SE=3. It appeared that the "hard" work for SE=3 was seen as a challenge, while for SE=0 hard work was a stumbling block. The 126 annotations for which DIFF=3, DOMORE=0 and the 77 annotations for which DIFF=3, DOMORE=1 were reread, for themes, and although this process did require making some inferences, the results were very clear.

Twenty-three of the 77(30%) annotations where the student indicated he or she wanted to do more hard work included some comment about learning more, getting better with practice, or value for future work. Twenty-four of the 77(31%) included comments about enjoying the work. Fourteen of the 77 (18%) included statements about challenge. None of the 77 annotations for DIFF= 3, DOMORE=1 indicated that the work was too hard.

Conversely, 8 of the 126 (6%) annotations where the student indicated he or she did not want to do more of the hard work included the words "too hard." Thirty-one of the 126 (25%) annotations simply pronounced the work hard, with no explanation. Twenty-two of the 126(17%) annotations included a statement about not enjoying the work. Only one (1%) mentioned a challenge.

The validity evidence from qualitative analyses of the students's own writings was very compelling, with one exception. Learning attitudes demonstrated by the annotations were consistent with the DIFF, DOMORE, and SE ratings. Key words and concepts were consistent with these ratings, too, with the exception of the 53 annotations whose student comments read they did not want to do more easy work because the work was "too easy", plus a few with similar viewpoints that did not use the phrase "too easy." The tone of these comments was confident and did not evince lack of self-efficacy for learning.

Thus the constructed measure of students' self-efficacy to accomplish their work did not stand up to a validity study. Data from the SE variable (DIFF X DOMORE) stood up to reliability checks and preliminary validity investigations. However, a content analysis of comments students wrote indicated that at least some of those students who did not want to do more of the kind of work on the work sample, even though they reported it was easy

work, felt that way because they considered the work too easy and therefore mastered. Wanting to move on to harder work because easier work is mastered is not consistent with a score of "0" on the SE scale.

Since the other categories of the SE scale did seem to match with student comments, especially noticeable in the comments of SE=3 students who reported enjoying handling academic challenge, the search for an SE measure from these annotations should not be abandoned. Portfolios present an opportunity for students to reflect on the achievement represented by particular work samples. The potential exists for a more specific measure of self- efficacy than heretofore available.

Study Two

Method

Data. Data were collected in 1995, during the entire second semester (January through May). Within each class in the evaluation project, portfolios from 6 students were collected according to a stratified (gifted, nonlabeled, and educational support) random sample. There were portfolio data from a total of 313 students, grades 1-10, in whose portfolios were included a total of 2862 annotated work samples. Portfolios were from mathematics (88%) and language arts (12%) classes. Fifty-two percent of students were male; 48% were female. Their status was 27% gifted, 47% nonlabeled, and 26% educational support. Data from student annotation forms and teacher ratings of achievement and time use were the same as for Study One, except that the SE variable (DIFF times DOMORE) was not calculated. The SE variable was dropped because of the results of the validity analyses in Study One.

Analyses. Two reliability analyses were done, to replicate those conducted for Study One. (1) Cronbach's alpha was calculated for DIFF and DOMORE ratings, treating annotations within one student's portfolio as items on the respective scales. Student was the unit of analysis. The rationale for this analysis was the same as for Study One, namely, that overall DIFF and DOMORE ratings for each student had been used in the evaluation and thus internal consistency was required. (2) Generalizability studies with a Person by Time design were conducted for DIFF and DOMORE scales. The rationale was again the same as for Study One, namely, to establish how many annotations were required for reliable overall scores and to determine whether repeated use of the same form compromised the validity of the information.

The external structure of DIFF and DOMORE ratings were again examined as validity evidence. A replication of the qualitative study of validity evidence, coding student written responses, was not possible within the time frame of the study.

Results and Discussion

Reliability. Internal consistency results for DIFF were slightly higher than those found in Study One. Alpha reliability ranged from .55 for 6 annotations per portfolio to .70 for 9 annotations (see Table 2.1, cf. .59 to .64 for Study One). Internal consistency results for DOMORE were slightly lower than those found in Study One. Reliability ranged from .65 for 7 annotations per portfolio to .72 for 10 annotations (see Table 2.1, cf. .66 to .80 for Study One). The number of students upon which these analyses were based was larger in Study Two than in Study One, and the estimates are therefore more stable.

Table 2.1 Reliability (Internal Consistency) of Scales for Six through Ten Annotations per Portfolio

Scale	k	α	n
DIFF	6	.55	274
	7	.63	240
	8	.67	207
	9	.70	185
	10	.66	157
DOMORE	6	.66	248
	7	.65	217
	8	.70	194
	9	.70	171
	10	.72	144

The generalizability results from Study Two were virtually identical to those from Study One (see Tables 2.2. and 2.3). The conclusion that repeated use of the same form over time did not have any effect was once again supported. Less than 1% of the variance in both DIFF and DOMORE was due to Time (Table 2.2). The sizes of the other variance components was similar to those in Study One. The finding that the Person-by-Time interaction was the largest effect for both DIFF and DOMORE was replicated. As in Study One, this large interaction term is more similar to results from performance assessments than from paper-and-pencil tests.

Both DIFF and DOMORE scales generalized at an acceptable level over repeated use in a portfolio (Table 2.3). As for Study One, relative and absolute generalizability coefficients were similar because the Time effect was almost zero. Generalizability coefficients for DIFF were virtually identical to those from Study One; generalizability coefficients for DOMORE were slightly lower than those for Study One (see Table 1.4).

Table 2.2 Variance Component Estimates for Scales, for Person x Time Design

Source of Variability	DIFF		DOMORE	
	Estimated Variance Component	Percent Total Variability	Estimated Variance Component	Percent Total Variability
Person	.0955	19%	.0416	21%
Time	.0017	< 1%	.0003	< 1%
P x T	.3987	80%	.1573	80%

DIFF variance estimates based on 240 persons, 7 times
 DOMORE variance estimates based on 216 persons, 7 times

Table 2.3 Generalizability of Scales, for Six through Ten Annotations per Portfolio

Scale	k	$\hat{\rho}^2$	$\hat{\phi}$
DIFF	6	.59	.59
	7	.63	.63
	8	.66	.66
	9	.68	.68
	10	.71	.70
DOMORE	6	.61	.61
	7	.65	.65
	8	.68	.68
	9	.70	.70
	10	.73	.72

DIFF estimates based on 240 persons, 7 times
 DOMORE estimates based on 216 persons, 7 times

Validity Table 2.4 shows that the correlational validity evidence from Study Two was in the same direction as for Study One but was stronger. Both DIFF and DOMORE were significantly but weakly related to the percent of instructional objectives attained (of those attempted for the portfolio, PCTACH) and to teacher' judgments of students time use (TIME) and evaluation of achievement quality (EVAL). Teachers judge more efficient time use for students who rated their work samples less difficult overall. Teacher judged more efficient time use for students who indicated more interest in doing their work. These findings accord with the significant results from Study One, but offer slightly stronger evidence that students' in-class behavior, use of lesson and study time, is consistent with their annotations. For Study Two, unlike Study One, there were also significant but weak relationships between the quality of students' work (percent achievement and evaluation) and the DIFF and DOMORE variables. More difficult ratings on annotations were associated with lower percentages of achievement of attempted objectives and with lower teacher evaluations of work quality. Willingness to do more of the same kind of work was associated with higher percentages of achievement of attempted objectives and with higher teacher evaluations of work quality (Table 2.4). Study Two's results may be more accurate because of the length of data collection . In 1995, the school district required that portfolios be kept for the entire second semester. In 1994 (Study One's data), the length of collection varied from a few weeks to an entire semester, depending on the class.

Table 2.4 Correlations between Scales and Selected External Measures

	PCTACH	TIME	EVAL
DIFF	- . 15 *	- . 26 **	- . 13 *
DOMORE	- . 14 *	- . 18 **	. 12 *

* p < .05, **p < .01, n ranged from 242 to 284 students

PCTACH = percent of learning objectives attempted within the portfolio that were achieved
 TIME = efficiency of time use, 0 = inefficient, 1 = efficient, average over all learning objectives represented in the portfolio

EVAL = teacher evaluation of achievement, on a scale of 1 to 4, averaged over all learning objectives represented in the portfolio

In Study One, lack of relationship between annotations and achievement was interpreted to mean that the annotations measured something different from achievement. This interpretation may still be the one that makes the most sense. Relationships between annotations and achievement were significant (and logical, since the annotations were affixed to school work samples meant to demonstrate achievement); nevertheless, they were small enough that they do not explain much of the variance in achievement. Clearly, additional information, besides achievement, is contained in the DIFF and DOMORE ratings.

Conclusion

The basic indicators on the Student Annotation Form, the DIFF and DOMORE questions, did stand up to reliability and validity study. Reliability and validity of the items on the Student Annotation Form were confirmed for the purpose of a districtwide evaluation, where efficient and aggregatable measures were required. An attempt to construct a multiplicative self-efficacy measure from these ratings did not stand up to a validity study. The Student Annotation Form provided a simple, useful tool for indicating student judgments about individual work samples in a portfolio. Aggregated over time, during the longitudinal process of keeping a portfolio, these annotations also provided simple but useful measures of students' critical judgments of their work output—their perceptions of its difficulty and their willingness to do more of it. These are outcomes of general interest in current educational practice.

References

- Azar, M., Cass, M. Cipielliewski, J., Jordan, R., Kraetke, S., & Markel, K. (1993). *Sit beside me...Macomb portfolio project. 1990-93*. Clinton Township, MI: Macomb Intermediate School District. (ERIC Document Reproduction Service No. ED 361 385)
- Brennan, R. L., Gao, X., & Colton, D.A. (1995). Generalizability analyses of WorkKeys listening and writing tests. *Educational and Psychological Measurement, 55*, 157-176.
- Calfee, R. C., & Perfumo, P. (1993). *Student portfolios and teacher logs: Blueprint for a revolution in assessment*. Berkeley, CA, & Pittsburgh, PA: Center for the Study of Writing. (ERIC Document Reproduction Service No. ED 362 887)
- Farnan, N., & Kelly, P. (1991). Keeping track: Creating assessment portfolios in reading and writing. *Reading, Writing, and Learning Disabilities, 7*, 255-269.
- Ferguson, S., & Maples, C. (1992). Portfolios: Windows on learning: Zeroing in on math abilities. *Learning, 21* (3), 38-41.

- Goodman, K. S., Bird, L. B., & Goodman, Y.M. (1992). *The whole language catalog supplement on authentic assessment*. Santa Rosa, CA: American School Publishers.
- Herman, J.L., & Winters, L. (1994). Portfolio research: A slim collection. *Educational Leadership*, 52 (2), 48 - 55.
- Klein, S.P., McCaffrey, D., Stecher, B., & Koretz, D. (1995). The reliability of mathematics portfolio scores: Lessons from the Vermont experience. *Applied Measurement in Education*, 8, 243 - 260.
- Roe, M., & Vukelich, C. (1995, April). *That was then and this is now: A longitudinal study of teachers's portfolio practices*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.
- Salomon, G. (1984). Television is "easy" and print is "tough": The differential investment of mental effort as a function of perceptions and attributions. *Journal of Educational Psychology*, 76. 647 - 658.
- Wang, M. C. (1992). *Adaptive education strategies: Building on diversity*. Baltimore: Paul H. Brooks Publishing.
- Zimmerman, B. J., & Martinez-Pons, M. (1992) Perceptions of efficacy and strategy use in the self-regulation of learning. In D.H Schunk & J.L. Meece (Eds.), *Student perceptions in the classroom* (pp. 185 - 207). Hillsdale, NJ: Lawrence Erlbaum.

The Development of Multi-level Causal Analysis Technique: the Application of HLM Program

Nikom Nak-ai

ABSTRACT

The purposes of this research were to develop the multi-level causal analysis technique by using the HLM program and to test such technique with empirical data. The results can be summarized as follow :

1. The multi-level analysis and path analysis can be applied to analyze the multi-level data via the multi-level causal analysis technique: HLM in path analysis developed by the researcher.

2. The result of using the multi-level causal analysis technique with the application of HLM program can be drawn from information and analysis process when empirical data was tested by the analysis technique developed by the researcher. They are as follows :

2.1 The multi-level causal analysis technique can be used to analyze the fixed effect, random effect or between - class variance by taking into consideration the significance of t-test and Chi - square test. According to HLM analysis, the simple model can be used to analyze the causal micro model and the hypothetical model to the macro model with regression analysis by SPSS/PC .

2.2 The multi-level causal analysis technique can be used to test the coherence of such models and empirical data in both causal micro model analysis and causal macro model analysis by Specht's method.

2.3 The multi-level causal analysis technique can be used to analyze the decomposition of correlation analysis between variables in direct effect, indirect effect, and total effect in both causal micro model analysis and causal macro model analysis.

* Master's thesis of Department of Educational Research, Chulalongkorn University under the advice of Assoc. Prof. Sirichai Karnjanawasee, Ph.D.

การพัฒนาเทคนิควิเคราะห์เชิงสาเหตุแบบพหุระดับ การประยุกต์ใช้โปรแกรมเอชแอลเอ็ม*

นิคม นาคอ้าย

บทคัดย่อ

การวิจัยนี้ มีจุดมุ่งหมายเพื่อพัฒนาเทคนิควิเคราะห์เชิงสาเหตุแบบพหุระดับโดยการประยุกต์ใช้โปรแกรมเอชแอลเอ็ม และทดลองใช้เทคนิควิเคราะห์เชิงสาเหตุแบบพหุระดับ โดยนำเทคนิคที่พัฒนาขึ้นไปตรวจสอบกับข้อมูลเชิงประจักษ์ ผลการวิจัยที่สำคัญมีดังนี้

1. *Multilevel Analysis* และ *Path Analysis* สามารถประยุกต์ใช้ในการวิเคราะห์ *Multilevel Data* ได้ โดยอาศัยเทคนิควิเคราะห์สาเหตุแบบพหุระดับ (*multilevel causal analysis technique*) ที่ผู้วิจัยพัฒนาขึ้นจากหลักการของการวิเคราะห์พหุระดับด้วยโปรแกรมเอชแอลเอ็มในการวิเคราะห์โครงสร้างความสัมพันธ์เชิงสาเหตุของตัวแปร ตามโมเดล *Path Analysis*

2. ผลการทดลองใช้เทคนิควิเคราะห์เชิงสาเหตุแบบพหุระดับโดยการประยุกต์ใช้โปรแกรมเอชแอลเอ็ม สามารถพิจารณาได้จากสารสนเทศและกระบวนการวิเคราะห์ เมื่อนำเทคนิควิเคราะห์ที่ผู้วิจัยพัฒนาขึ้นไปตรวจสอบกับข้อมูลเชิงประจักษ์ ดังนี้

2.1 เทคนิควิเคราะห์เชิงสาเหตุแบบพหุระดับ สามารถวิเคราะห์อิทธิพลคงที่ อิทธิพลสุ่มหรือความแปรปรวนระหว่างห้องเรียน โดยการพิจารณาความมีนัยสำคัญของการทดสอบ t และการทดสอบ χ^2 ตามรูปแบบของการวิเคราะห์ในโมเดลเอชแอลเอ็ม ทั้งโมเดลเชิงสาเหตุระดับนักเรียน ซึ่งใช้การวิเคราะห์ขั้น *Simple Model* และโมเดลเชิงสาเหตุระดับชั้นเรียนในการวิเคราะห์ขั้น *Hypothetical Model* ประกอบการวิเคราะห์ที่ถดถอยด้วยโปรแกรม *SPSS/PC**

2.2 เทคนิควิเคราะห์เชิงสาเหตุแบบพหุระดับสามารถทดสอบความสอดคล้องของโมเดลกับข้อมูลเชิงประจักษ์ได้ ทั้งในโมเดลเชิงสาเหตุระดับนักเรียน และโมเดลเชิงสาเหตุระดับชั้นเรียนด้วยวิธีของสเปค

2.3 เทคนิควิเคราะห์เชิงสาเหตุแบบพหุระดับ สามารถวิเคราะห์แยกค่าสหสัมพันธ์ระหว่างตัวแปร เพื่อศึกษาผลกระทบทางตรง ผลกระทบทางอ้อม และผลกระทบรวมได้ทั้งการวิเคราะห์โมเดลเชิงสาเหตุระดับนักเรียน และโมเดลเชิงสาเหตุระดับชั้นเรียน

* วิทยานิพนธ์ครุศาสตรมหาบัณฑิต ปีการศึกษา 2539 โดยมีรองศาสตราจารย์ ดร. ศิริชัย กาญจนวาสี เป็นอาจารย์ที่ปรึกษา

ความเป็นมาและความสำคัญของปัญหา

ข้อมูลทางการศึกษามีลักษณะที่สำคัญประการหนึ่ง คือการสอดแทรก (nested) และเป็นระดับลดหลั่นกัน (hierarchy) อยู่ในรูปของข้อมูลพหุระดับ (multilevel data) อาทิเช่น นักเรียนในระดับเดียวกันถูกจัดกลุ่มเข้าด้วยกันเพื่อเป็นห้องเรียน ห้องเรียนหลาย ๆ ห้องเรียนรวมกันเป็นระดับชั้น ระดับชั้นหลาย ๆ ระดับชั้นรวมกันเป็นระดับการศึกษา ระดับการศึกษาหลาย ๆ ระดับ การศึกษารวมกันอยู่ในโรงเรียน เป็นต้น เมื่อพิจารณาถึงบุคลากรทางการศึกษาจะเห็นได้ว่าการจัดการบริหารงานเป็นระดับชั้นเพื่อทำหน้าที่ต่าง ๆ กัน เช่น ผู้อำนวยการหรือครูใหญ่ ทำหน้าที่บริหารโรงเรียน ครูมีหน้าที่รับผิดชอบการเรียนการสอนในชั้นเรียน ส่วนนักเรียนซึ่งรวมกันเป็นกลุ่มมีหน้าที่ศึกษาเล่าเรียนในชั้นใดชั้นหนึ่ง เป็นต้น ดังนั้นธรรมชาติของตัวแปรทางการศึกษาจึงมีลักษณะการจัดเป็นระดับชั้นไปโดยปริยาย ดังนั้นหากนักวิจัยสนใจวิเคราะห์อิทธิพลของตัวแปรอิสระต่างระดับที่มีต่อผลสัมฤทธิ์ทางการเรียน ซึ่งเป็นตัวแปรระดับนักเรียนในรูปแบบของการวิเคราะห์รวมเสมือนหนึ่งว่าตัวแปรอิสระเหล่านั้นอยู่ในระดับเดียวกันโดยใช้เทคนิควิเคราะห์ถดถอยพหุคูณ (multiple regression analysis) ซึ่งเป็นการละเลยต่อโครงสร้างของระดับข้อมูล ผลที่ตามมาคือเกิดความผิดพลาดในการสรุปผลระหว่างระดับ (aggregation bias) เนื่องจากมีความผิดพลาดในการประมาณค่าสัมประสิทธิ์การถ่วงน้ำหนัก ตลอดจนความคลาดเคลื่อนในการทำนายมีความแปรปรวนสูงไม่คงที่ นอกจากนี้แล้วการวิเคราะห์รวมโดยไม่สนใจความแตกต่างระหว่างหน่วยวิเคราะห์จะไม่สามารถคำนวณค่าความแปรปรวนภายในหน่วยหรือในกลุ่ม (within group variance) อันเป็นการละเลยการศึกษาปฏิสัมพันธ์ของตัวแปรที่อยู่ต่างระดับกัน (Raudenbush and Bryk, 1986; Kanjanawasee, 1989) จากปัญหาดังกล่าวข้างต้นทำให้นักวิจัยทางการศึกษาหลายท่าน ได้เสนอเทคนิควิธีการออกแบบการวิจัย และการวิเคราะห์พหุระดับ (multilevel analysis) ซึ่งมีความเหมาะสมกับโครงสร้างและข้อมูลทางการศึกษา (Cronbach, 1976; Burstein, Lin and Capell, 1978; Burstein, 1980a, 1980b; Mason, Wong and Entwistle, 1984; Aitkin and Longford, 1986; Raudenbush and Bryk, 1986; Raudenbush, Chinapat and Mohamed, 1992)

การวิเคราะห์พหุระดับจึงเป็นเทคนิควิธีการทางสถิติที่ใช้วิเคราะห์ข้อมูลที่มีตัวแปรอิสระหลายตัว และตัวแปรอิสระเหล่านั้นสามารถจัดเป็นระดับได้อย่างน้อยสองระดับขึ้นไป โดยตัวแปรระดับเดียวกันต่างมีความสัมพันธ์ซึ่งกันและกันและได้รับอิทธิพลร่วมกันจากตัวแปรระดับที่สูงกว่า โดยความพยายามในการพัฒนาวิธีการวิเคราะห์ข้อมูลพหุระดับ ได้ดำเนินมาอย่างต่อเนื่อง ดังจะเห็นได้จากการเปรียบเทียบโปรแกรมคอมพิวเตอร์ 4 โปรแกรม โดย Kreft and de Leeuw (1995) โปรแกรม ดังกล่าวประกอบด้วย VARCL ของ Longford (1988), ML/3 ของ Goldstein

(1989), HLM ของ Raudenbush and Bryk (1989) และ GENMOD ของ Mason (1988) ซึ่ง Kreft and de Leeuw (1995) ได้ให้ข้อสรุปโดยทั่วไปเกี่ยวกับประสิทธิภาพของโปรแกรม พอสรุปได้ดังนี้

$$\text{HLM} \geq \text{VARCL} \geq \text{ML}/3 \geq \text{XLISP}$$

ประสิทธิภาพของโปรแกรมที่นำเสนอข้างต้นสรุปได้ว่า HLM มีกระบวนการวิเคราะห์ที่ง่ายกว่า VARCL แต่ก็ให้ผลการวิเคราะห์ออกมาเหมือนกัน ถ้าโมเดลมีเงื่อนไขที่เหมาะสม ผลการวิเคราะห์เปรียบเทียบระหว่าง VARCL กับ ML/3 จะไม่แตกต่างกันมากนัก ในขณะที่ ML/3 จะให้สารสนเทศบางประการเกี่ยวกับค่าสถิติที่มีความเป็นทั่วไปมากกว่า HLM เช่นค่าเศษเหลือ (Residual) เป็นต้น อย่างไรก็ตาม HLM Approach นับว่าถูกพัฒนามาจากพื้นฐานของโมเดลที่ถูกต้องกว่าแบบอื่นทั้งหมด ทั้งนี้การเรียงลำดับประสิทธิภาพของโปรแกรมนั้นขึ้นอยู่กับวัตถุประสงค์ที่แตกต่างกันของผู้ใช้โปรแกรมเป็นสำคัญ

อย่างไรก็ตาม แนวคิดในการวิเคราะห์ข้อมูลพหุระดับด้วยเทคนิคเอชแอลเอ็มที่กล่าวมาข้างต้นนั้น เป็นการวิเคราะห์ข้อมูลเพื่ออธิบายอิทธิพลของชุดของตัวแปรต้นที่มีต่อตัวแปรตามด้วยสมการพยากรณ์ (regression equation) เท่านั้น หรือกล่าวได้ว่า การวิเคราะห์ข้อมูลพหุระดับที่ผ่านมา ยังมีได้ทำการวิเคราะห์ในเชิงสาเหตุที่ตัวแปรต้นมีต่อตัวแปรตามในรูปสมการพยากรณ์ที่แสดงความเกี่ยวข้องสัมพันธ์กันทั้งหมดของตัวแปร ในลักษณะของ Path Model นั้นหมายความว่า การวิเคราะห์ข้อมูลพหุระดับด้วยเทคนิคเอชแอลเอ็ม มุ่งวิเคราะห์ตัวแปรพหุระดับ โดยศึกษาอิทธิพลของตัวแปรต้นที่มีต่อตัวแปรตามในแต่ละระดับแต่เพียงอย่างเดียว แต่ยังมีได้มุ่งศึกษาความสัมพันธ์เชิงสาเหตุระหว่างตัวแปรในโมเดลการวิจัยแต่ละระดับ ดังนั้นจึงน่าจะได้ขยายขอบเขตของการวิเคราะห์ตัวแปรพหุระดับ ให้สามารถศึกษาถึงความเป็นสาเหตุระหว่างตัวแปร ตามเทคนิคของการวิเคราะห์เชิงสาเหตุ (path analysis) ด้วย

การนำแนวคิดของการวิเคราะห์ตัวแปรพหุระดับ มาทำการวิเคราะห์ในโมเดล Path Analysis ในรูปของการวิเคราะห์เชิงสาเหตุแบบพหุระดับ (multilevel causal analysis) จึงเป็นการพัฒนาวิธีวิทยาทางการวิจัยแบบใหม่ ซึ่งยังไม่มีผู้ใดพัฒนามาก่อน และเป็นแนวทางหนึ่งของการพัฒนาเทคนิควิธีทางสถิติที่มุ่งศึกษาถึงความสัมพันธ์เชิงสาเหตุระหว่างตัวแปร นอกเหนือจากการพิจารณาถึงโครงสร้างของระดับข้อมูลแต่เพียงอย่างเดียว ผู้วิจัยจึงมีความสนใจที่จะศึกษาว่า การวิเคราะห์ความสัมพันธ์เชิงสาเหตุแบบพหุระดับโดยการประยุกต์ใช้โปรแกรมเอชแอลเอ็ม ซึ่งได้รับการยอมรับว่ามีประสิทธิภาพสูงสุดในการวิเคราะห์พหุระดับนั้น ควรมีหลักการ วิธีการและกระบวนการในการวิเคราะห์อย่างไร สารสนเทศที่ได้มีลักษณะเช่นใด และควรมีแนวทางในการแปล

ความหมายอย่างไร ทั้งนี้เพื่อให้ผลการวิเคราะห์มีประสิทธิภาพสูงสุดในแง่ของการอธิบายความสัมพันธ์ระหว่างตัวแปรในแต่ละระดับ ที่ส่งผลต่อตัวแปรตามในโมเดลการวิจัยต่อไป

วัตถุประสงค์ของการวิจัย

1. เพื่อพัฒนาเทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับโดยการประยุกต์ใช้โปรแกรมเอชแอลเอ็ม
2. เพื่อทดลองใช้เทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับ โดยนำเทคนิควิธีที่พัฒนาขึ้นไปตรวจสอบกับข้อมูลเชิงประจักษ์

สมมติฐานของการวิจัย

จากแนวคิดเบื้องต้นที่ Jan de Leeuw (อ้างใน Raudenbush and Bryk, 1992) ได้กล่าวว่า หากนักวิจัยขยายรูปแบบการวิเคราะห์ออกไปจนสามารถสร้างโมเดลการวิเคราะห์เชิงสาเหตุแบบพหุระดับได้อย่างสมเหตุสมผลแล้ว ย่อมคาดหวังถึงผลการวิเคราะห์ที่มีประสิทธิภาพและมีความแม่นยำมากขึ้น ผู้วิจัยจึงตั้งสมมติฐานการวิจัยไว้ดังนี้

1. Multilevel Analysis และ Path Analysis นำจะประยุกต์ใช้ในการวิเคราะห์ Multilevel Data ได้ โดยอาศัยเทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับที่ผู้วิจัยพัฒนาขึ้นจากหลักการของการวิเคราะห์พหุระดับด้วยโปรแกรมเอชแอลเอ็มในการวิเคราะห์โครงสร้างความสัมพันธ์เชิงสาเหตุของตัวแปร ตามโมเดล Path Analysis

2. ผลการทดลองใช้เทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับโดยการประยุกต์ใช้โปรแกรมเอชแอลเอ็ม สามารถพิจารณาได้จากสารสนเทศและกระบวนการวิเคราะห์ เมื่อนำเทคนิควิธีวิเคราะห์ที่ผู้วิจัยพัฒนาขึ้นไปตรวจสอบกับข้อมูลเชิงประจักษ์ ดังนี้

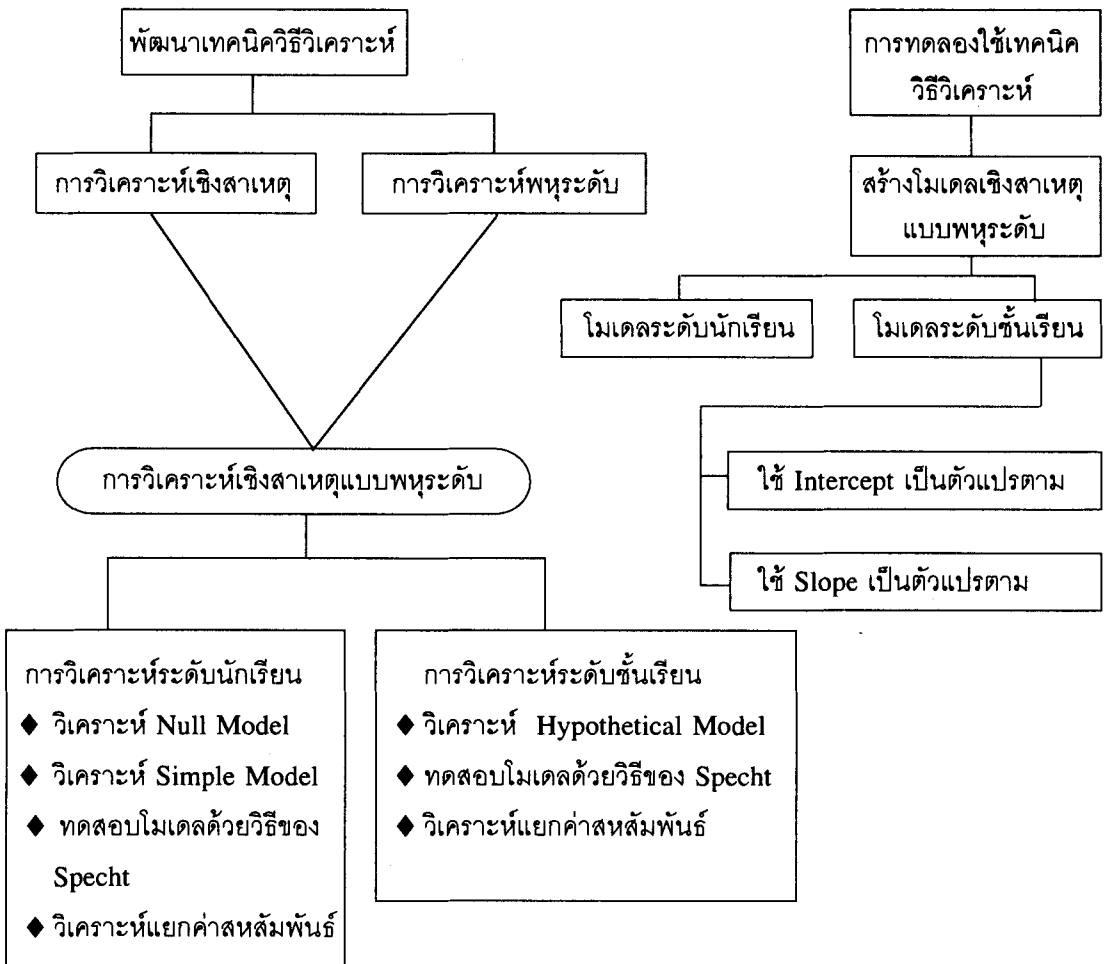
- 2.1 เทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับ สามารถวิเคราะห์อิทธิพลคงที่ (fixed effects) อิทธิพลสุ่ม (random effects) หรือความแปรปรวนระหว่างห้องเรียน (between-class variance) โดยการพิจารณาความมีนัยสำคัญของการทดสอบ t และการทดสอบ χ^2 ตามรูปแบบของการวิเคราะห์ในโมเดลเอชแอลเอ็ม

- 2.2 เทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับสามารถทดสอบความสอดคล้องของโมเดลกับข้อมูลเชิงประจักษ์ได้ทั้งในโมเดลการวิเคราะห์ระดับนักเรียนและระดับชั้นเรียน

- 2.3 เทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับ สามารถวิเคราะห์แยกค่าสหสัมพันธ์ระหว่างตัวแปร เพื่อศึกษาผลกระทบทางตรง (direct effects) ผลกระทบทางอ้อม (indirect effects) และผลกระทบรวม (total effects) ได้ทั้งการวิเคราะห์ระดับนักเรียน และระดับชั้นเรียน

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้ดำเนินการโดยใช้ระเบียบวิธีวิจัยเชิงพัฒนา มีวัตถุประสงค์เพื่อพัฒนาเทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับ (multilevel causal analysis technique) โดยการประยุกต์ใช้โปรแกรมเอชแอลเอ็มในการวิเคราะห์โครงสร้างความสัมพันธ์เชิงสาเหตุระหว่างตัวแปรต่างระดับ เทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับที่ผู้วิจัยพัฒนาขึ้นนี้ จะนำไปทดลองใช้กับข้อมูลเชิงประจักษ์อันเป็นข้อมูลการวิจัยของประเสริฐ เตชะนาราเกียรติ (2532) มีขั้นตอนการดำเนินการวิจัยดังแผนภาพ ดังต่อไปนี้



ภาพที่ 1 แสดงขั้นตอนการดำเนินการดำเนินการวิจัย

1. การพัฒนาเทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับ

1.1 ศึกษาค้นคว้าเอกสาร ตำรา และงานวิจัยที่เกี่ยวข้องกับการวิเคราะห์ เชิงสาเหตุ (path analysis)

1.2 ศึกษาค้นคว้าเอกสารตำราและงานวิจัยที่เกี่ยวข้องกับการวิเคราะห์พหุระดับ (multilevel analysis)

1.3 พัฒนาเทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับ (multilevel causal analysis) โดยประยุกต์วิธีการวิเคราะห์ถดถอยเป็นลำดับขั้นตามโมเดลเชิงสาเหตุ กับเทคนิคการวิเคราะห์พหุระดับด้วยโปรแกรมเอชแอลเอ็ม ดังนี้

1.3.1 พัฒนาการวิเคราะห์ขั้น Null Model เพื่อให้สามารถพิจารณาความผันแปรของตัวแปรตามของโมเดลเชิงสาเหตุระดับนักเรียนได้ทุกตัว ในการวิเคราะห์ถดถอยแบบเป็นลำดับขั้นตอนด้วยโปรแกรมเอชแอลเอ็ม

1.3.2 พัฒนาการวิเคราะห์ขั้น Simple Model เพื่อให้สามารถพิจารณาอิทธิพลคงที่ อิทธิพลสุ่มหรือความแปรปรวนระหว่างห้องเรียน ในการวิเคราะห์โมเดลเชิงสาเหตุระดับนักเรียน (causal micro model) ด้วยการวิเคราะห์ถดถอยแบบเป็นลำดับขั้นตอนด้วยโปรแกรมเอชแอลเอ็ม ทั้งโมเดลเชิงสาเหตุระดับนักเรียนแบบเต็มรูปและโมเดลเชิงสาเหตุระดับนักเรียนตามสมมติฐาน

1.3.3 พัฒนารูปแบบการทดสอบความสอดคล้องของโมเดลเชิงสาเหตุระดับนักเรียนตามสมมติฐาน กับข้อมูลเชิงประจักษ์ ด้วยวิธีของสเปค

1.3.4 พัฒนารูปแบบการวิเคราะห์แยกค่าสหสัมพันธ์ระหว่างตัวแปร ในโมเดลเชิงสาเหตุระดับนักเรียน ที่ผ่านการทดสอบความสอดคล้องกับข้อมูลเชิงประจักษ์แล้ว

1.3.5 พัฒนาการวิเคราะห์ขั้น Hypothetical Model เพื่อให้สามารถพิจารณาอิทธิพลคงที่ อิทธิพลสุ่มหรือความแปรปรวนระหว่างห้องเรียน ในการวิเคราะห์โมเดลเชิงสาเหตุระดับชั้นเรียน (causal macro model) ด้วยการวิเคราะห์ถดถอยแบบเป็นลำดับขั้นตอนด้วยโปรแกรมเอชแอลเอ็ม ทั้งโมเดลเชิงสาเหตุระดับชั้นเรียนแบบเต็มรูปและโมเดลเชิงสาเหตุระดับชั้นเรียนตามสมมติฐาน

1.3.6 พัฒนารูปแบบการทดสอบความสอดคล้องของโมเดลเชิงสาเหตุระดับชั้นเรียนตามสมมติฐาน กับข้อมูลเชิงประจักษ์ ด้วยวิธีของสเปค

1.3.7 พัฒนารูปแบบการวิเคราะห์แยกค่าสหสัมพันธ์ระหว่างตัวแปร ในโมเดลเชิงสาเหตุระดับชั้นเรียน ที่ผ่านการทดสอบความสอดคล้องกับข้อมูลเชิงประจักษ์แล้ว

1.4 สร้างโมเดลเชิงสาเหตุแบบพหุระดับ โดยการศึกษาค้นคว้าเอกสาร ตำราและงานวิจัยที่เกี่ยวข้อง เพื่อทดลองใช้เทคนิควิธีวิเคราะห์ที่พัฒนาขึ้น

2. การทดลองใช้เทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับ

2.1 นำข้อมูลงานวิจัยของประเสริฐ เตชะนาราเกียรติ (2532) ซึ่งบันทึกข้อมูลในรูปแบบของโปรแกรม Q-EDIT ในจำนวนที่ชนิดก่อนขนาด 1.44 เมกกะไบต์ มาทำการคัดเลือกตัวแปรตามโมเดลเชิงสาเหตุที่ใช้ในการวิจัย จากนั้นจึงแบ่งข้อมูลที่คัดเลือกออกเป็น 2 ระดับตามรูปแบบของฟอร์แทนฟอร์มเมต คือ ระดับนักเรียนและระดับชั้นเรียน เพื่อใช้ในการวิเคราะห์ด้วยโปรแกรมเอสแอลเอ็ม โดยตัวแปรระดับนักเรียนประกอบด้วย ผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ แรงจูงใจใฝ่สัมฤทธิ์ เจตคติต่อวิชาคณิตศาสตร์ เชาวน์ปัญญา และรายได้ของผู้ปกครอง ตัวแปรระดับชั้นเรียนประกอบด้วย คุณภาพการสอนของครู ความเป็นผู้นำทางการวิชาการของผู้บริหารโรงเรียน และขนาดของโรงเรียน

2.2 จัดเตรียม SYSTEM FILES ชื่อ MICRO.SSM จากแฟ้มข้อมูลทั้ง 2 ระดับที่เตรียมไว้ เพื่อนำไปใช้วิเคราะห์ต่อไป

2.3 ทำการวิเคราะห์ถดถอยด้วยโปรแกรมเอสแอลเอ็ม ในโมเดลเชิงสาเหตุแบบพหุระดับซึ่งแบ่งออกเป็นระดับนักเรียน และระดับชั้นเรียน

ผลการพัฒนาและผลการทดลองใช้เทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับ

ผลการพัฒนาและผลการทดลองใช้เทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับ ผู้วิจัยได้นำเสนอ โดยแบ่งออกเป็น 2 ตอน ดังนี้

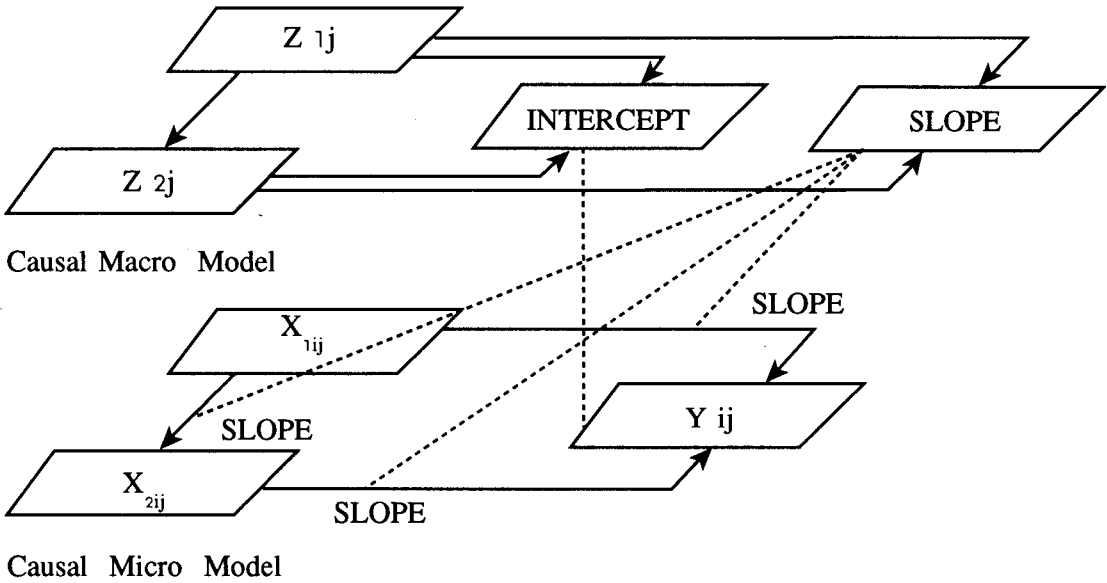
ตอนที่ 1 ผลการพัฒนาเทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับ โดยการประยุกต์ใช้โปรแกรมเอสแอลเอ็ม

จากการศึกษาค้นคว้าเอกสาร ตำรา และงานวิจัยที่เกี่ยวข้องกับการวิเคราะห์เชิงสาเหตุ (path analysis) และการวิเคราะห์พหุระดับ (multilevel analysis) ในเชิงของหลักการและกระบวนการวิเคราะห์ พบว่า สามารถพัฒนาเทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับ (multilevel causal analysis) โดยประยุกต์วิธีการวิเคราะห์ถดถอยเป็นลำดับขั้นตามโมเดลเชิงสาเหตุ กับเทคนิคการวิเคราะห์พหุระดับด้วยโปรแกรมเอสแอลเอ็ม ดังนี้

1.1 เทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับ สามารถทำการวิเคราะห์ ชั้น Null Model เพื่อให้สามารถพิจารณาความผันแปรของตัวแปรตามของโมเดลเชิงสาเหตุระดับนักเรียนได้ทุกตัว ในการวิเคราะห์ถดถอยแบบเป็นลำดับขั้นตอนด้วยโปรแกรมเอสแอลเอ็ม

1.2 เทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับสามารถทำการวิเคราะห์ ชั้น Simple Model และ Hypothetical Model เพื่อให้สามารถพิจารณาอิทธิพลของ อิทธิพลสุ่มหรือความแปรปรวนระหว่างห้องเรียน ในการวิเคราะห์โมเดลเชิงสาเหตุระดับนักเรียน (causal micro model) และ

โมเดลเชิงสาเหตุระดับชั้นเรียน (causal macro model) ด้วยการวิเคราะห์ถดถอยแบบเป็นลำดับขั้นตอนด้วยโปรแกรมเอชแอลเอ็ม ทั้งโมเดลแบบเต็มรูปและโมเดลตามสมมติฐาน ผู้วิจัยจึงขอนำเสนอโมเดลเบื้องต้นของการวิเคราะห์เชิงสาเหตุแบบพหุระดับที่พัฒนาขึ้น โดยสมมุติว่าตัวแปรทั้งหมดในโมเดลเป็นตัวแปรที่สังเกตได้โดยตรง (manifest variables model) และอยู่ในรูปของโมเดลที่ไม่มีความสัมพันธ์ย้อนกลับ (recursive model) ดังนี้

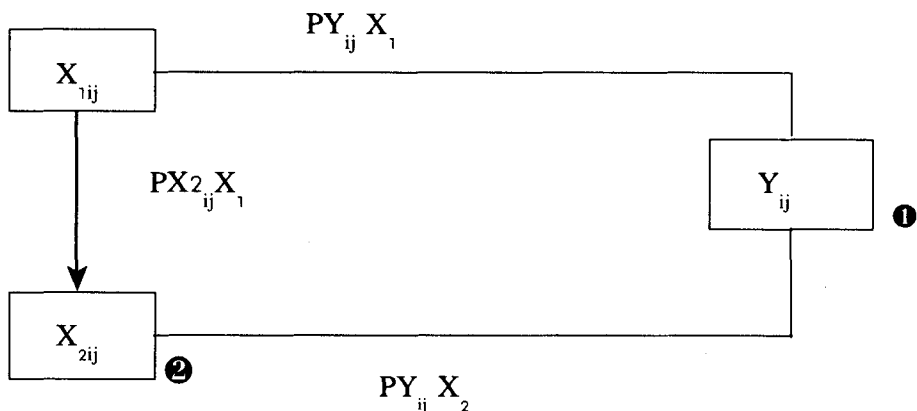


ภาพที่ 2 แสดงโมเดลรวมของการวิเคราะห์

- เมื่อ Y_{ij} เป็นตัวแปรตาม เช่นผลสัมฤทธิ์ทางการเรียนของนักเรียนแต่ละคนในห้องที่ j
- X_{1ij} เป็นตัวแปรอิสระตัวที่ 1 ในระดับนักเรียน (causal micro level) ของนักเรียนแต่ละคน ในห้องที่ j
- X_{2ij} เป็นตัวแปรอิสระตัวที่ 2 ในระดับนักเรียน (causal micro level) ของนักเรียนแต่ละคน ในห้องที่ j
- Z_{1j} เป็นตัวแปรอิสระตัวที่ 1 ในระดับชั้นเรียน (causal macro level) ในห้องที่ j
- Z_{2j} เป็นตัวแปรอิสระตัวที่ 2 ในระดับชั้นเรียน (causal macro level) ในห้องที่ j

จากโมเดลข้างต้นมีข้อมูลแบ่งเป็นสองระดับ ระดับแรกเป็นข้อมูลระดับนักเรียน (causal micro level) ระดับที่สอง เป็นข้อมูลระดับชั้นเรียน (causal macro level) หากยึดแนวคิดของการวิเคราะห์พหุระดับ มาเป็นหลักในการวิเคราะห์ จะสามารถนำเสนอผลการพัฒนาโมเดลตลอดจนแนวทางการประยุกต์ใช้โปรแกรมเอชแอลเอ็ม มาวิเคราะห์ในระดับที่หนึ่ง หรือระดับนักเรียน เป็นเบื้องต้นก่อน ได้ดังนี้

1.2.1 ผลการพัฒนาการวิเคราะห์เชิงสาเหตุแบบพหุระดับ ในระดับนักเรียน (causal micro model) การวิเคราะห์ในขั้นตอนนี้ สามารถประยุกต์วิธีการวิเคราะห์ที่ดัดดอายเป็นลำดับขั้นตามโมเดลเชิงสาเหตุ กับเทคนิคการวิเคราะห์พหุระดับด้วยโปรแกรมเอชแอลเอ็ม ในขั้นตอนของ Simple Model ดังนี้



ภาพที่ 3 แสดง Causal Micro Model : กรณี 3 ตัวแปร

- ❶ ตัวแปรตามในการวิเคราะห์ที่ดัดดอด้วยเอชแอลเอ็มครั้งที่ 1
- ❷ ตัวแปรตามในการวิเคราะห์ที่ดัดดอด้วยเอชแอลเอ็มครั้งที่ 2

จากภาพ แสดงให้เห็นว่า โมเดลการวิเคราะห์ระดับที่หนึ่ง (causal micro model) สามารถอธิบายได้ในรูปของการวิเคราะห์เส้นทางความสัมพันธ์เชิงสาเหตุระหว่างตัวแปร โดยตัวแปร Y_{ij} ได้รับอิทธิพลทางตรงจากทั้งตัวแปร X_{1ij} และ X_{2ij} ขณะเดียวกันก็ได้รับอิทธิพลทางอ้อมจากตัวแปร X_{1ij} ซึ่งส่งผ่านตัวแปร X_{2ij} อีกด้วย การวิเคราะห์โมเดลระดับที่หนึ่งข้างต้นจะทำให้ทราบค่าสัมประสิทธิ์เส้นทางความสัมพันธ์เชิงสาเหตุ (path coefficient) ของตัวแปรต้นที่มีต่อตัวแปรตามทั้งอิทธิพลทางตรงและทางอ้อม ในระดับนักเรียน คือค่า $PY_{ij} X_1$, $PY_{ij} X_2$, $PX_{2ij} X_1$ และ ซึ่งแต่ละค่าจะมีอยู่ทุกห้องเรียนหมายความว่า ถ้าเก็บข้อมูลมา j ห้องก็จะได้แต่ละค่าเหล่านี้ j ค่า จากนั้นใช้ค่าดังกล่าวเป็นตัวแปรตามในการวิเคราะห์ระดับที่สองต่อไป ทั้งนี้หากยึดแนวทางการวิเคราะห์ด้วยโปรแกรมเอชแอลเอ็มแล้ว จะสามารถวิเคราะห์เชิงสาเหตุแบบพหุระดับในโมเดลการวิเคราะห์ระดับนักเรียนได้ โดยการแบ่งผลพารามิเตอร์ที่ต้องประมาณเป็นอิทธิพลคงที่ และอิทธิพลสุ่มหรือความแปรปรวนระหว่างห้องเรียน ดังนี้

1.) การวิเคราะห์อิทธิพลคงที่ (fixed effect)

เป็นการวิเคราะห์เพื่อให้ทราบว่า ตัวแปรอิสระ (X_{1ij} , X_{2ij}) และค่าคงที่ (intercept) ซึ่งเป็นค่าเฉลี่ยของตัวแปรตาม (Y_j) มีอิทธิพลต่อตัวแปรตามในการวิเคราะห์ (Y_{ij}) อย่างมีนัยสำคัญทางสถิติหรือไม่ โดยเอชแอลเอ็มจะใช้ t-test ทดสอบสมมติฐานเกี่ยวกับอิทธิพลดังกล่าวจากทุกหน่วยการวิเคราะห์ว่ามีค่าเป็นศูนย์หรือไม่ ถ้าไม่เป็นศูนย์แสดงว่าตัวแปรอิสระ (X_{1ij} , X_{2ij}) และค่าคงที่ (intercept) มีอิทธิพลต่อตัวแปรตามในการวิเคราะห์ทดโดยด้วยเอชแอลเอ็มครั้งที่ 1 (Y_{ij}) อย่างมีนัยสำคัญทางสถิติ ผลการวิเคราะห์ที่ได้ จะเป็นตัวบ่งชี้ความมีนัยสำคัญทางสถิติของเส้นทางความสัมพันธ์เชิงสาเหตุ ตามโมเดลการวิเคราะห์ระดับนักเรียนที่นำเสนอ ทั้งนี้ การวิเคราะห์โดยใช้ X_{2ij} เป็นตัวแปรตาม และใช้ X_{1ij} เป็นตัวแปรอิสระในการวิเคราะห์การถดถอยด้วยเอชแอลเอ็มลำดับถัดไป (ครั้งที่ 2) เพื่อให้สามารถพิจารณาเส้นทางความสัมพันธ์เชิงสาเหตุได้ครบถ้วนตามโมเดล ก็สามารถทำการวิเคราะห์ได้ในทำนองเดียวกันกับการวิเคราะห์โดยมี Y_{ij} เป็นตัวแปรตาม

ทั้งนี้หากพบว่าเส้นทางความสัมพันธ์เชิงสาเหตุใดที่อิทธิพลคงที่ (fixed effect) ไม่มีนัยสำคัญทางสถิติ ผู้วิจัยอาจพิจารณาตัดเส้นทางความสัมพันธ์เชิงสาเหตุนั้นออกจากโมเดล ผลการวิเคราะห์ที่ได้นอกจากจะทำให้ทราบถึงอิทธิพลคงที่ของตัวแปรอิสระที่มีต่อตัวแปรตามแต่ละตัว (Y_{ij} , X_{2ij}) ตามโมเดลเชิงสาเหตุระดับนักเรียนแล้ว ยังสามารถคำนวณค่า R^2 (coefficient of determination) หรือค่าประสิทธิผลการพยากรณ์ ซึ่งสามารถนำไปวิเคราะห์ความสอดคล้องโมเดลกับข้อมูลเชิงประจักษ์ ด้วยวิธีการของสเปค ตลอดจนสามารถนำค่าสัมประสิทธิ์การถดถอยมาตรฐานที่ได้ไปวิเคราะห์แยกค่าสหสัมพันธ์ (decomposition of correlation) เพื่อพิจารณาปริมาณผลกระทบทางตรง (direct effect) ผลทางอ้อม (indirect effect) และผลกระทบรวม (total effect) ของตัวแปรในโมเดลการวิเคราะห์ระดับนักเรียนได้

2.) การวิเคราะห์อิทธิพลสุ่ม (random effect)

เป็นการวิเคราะห์เพื่อให้ทราบว่าค่าคงที่ (intercept) และสัมประสิทธิ์การถดถอยที่ได้จากการวิเคราะห์อิทธิพลคงที่ในโมเดลเชิงสาเหตุระดับนักเรียน มีความผันแปร (vary) ระหว่างห้องเรียนอย่างมีนัยสำคัญทางสถิติหรือไม่ โดยเอชแอลเอ็มจะใช้ χ^2 ทดสอบความแปรปรวนของค่าคงที่และสัมประสิทธิ์การถดถอยดังกล่าว ว่ามีค่าเป็นศูนย์หรือไม่ ถ้าไม่เป็นศูนย์ หรือค่า χ^2 มีนัยสำคัญทางสถิติ ก็แสดงว่าค่าคงที่และสัมประสิทธิ์การถดถอย มีความผันแปรระหว่างห้องเรียนเพียงพอที่จะวิเคราะห์หาตัวแปรอิสระในระดับชั้นเรียนมาอธิบายความผันแปรดังกล่าว หรือหมายความว่า เส้นทางความสัมพันธ์เชิงสาเหตุ (หรือสัมประสิทธิ์การถดถอยแต่ละค่า) และค่าเฉลี่ยของตัวแปรตาม ที่มีนัยสำคัญทางสถิติจากการวิเคราะห์อิทธิพลสุ่ม น่าจะได้รับอิทธิพลจากตัวแปร

อิสระระดับชั้นเรียน แต่หาก χ^2 ที่ได้ไม่มีนัยสำคัญทางสถิติ ก็แสดงว่าค่าคงที่และสัมประสิทธิ์การถดถอย ไม่มีความผันแปรระหว่างห้องเรียนเพียงพอที่จะวิเคราะห์หาตัวแปรอิสระในระดับชั้นเรียน มาอธิบายความผันแปรดังกล่าว และสามารถตั้งให้ข้อจำกัดในการวิเคราะห์ได้ โดยการควบคุมให้มีค่าเป็นศูนย์

ประเด็นสำคัญประการหนึ่ง ที่ต้องพิจารณา คือ การวิเคราะห์อิทธิพลคงที่และอิทธิพลสุ่มดังกล่าวข้างต้น จะต้องทำการวิเคราะห์ทั้งโมเดลเชิงสาเหตุแบบเต็มรูป (full model) และโมเดลเชิงสาเหตุตามสมมติฐาน (proposed model) ทั้งนี้เพื่อให้ได้โมเดลเชิงสาเหตุสุดท้ายหรือโมเดลแต่งใหม่ (trimmed model) ที่มีลักษณะประหยัด (parsimony) ก่อนที่จะทำการทดสอบความสอดคล้องของโมเดลกับข้อมูลเชิงประจักษ์ต่อไป

3.) ผลการพัฒนากการทดสอบความสอดคล้องของโมเดลเชิงสาเหตุระดับนักเรียนตามสมมติฐานกับข้อมูลเชิงประจักษ์ด้วยวิธีของสเปค พบว่า เนื่องจากโปรแกรมเอชแอลเอ็ม มุ่งวิเคราะห์พหุระดับด้วยหลักการของการวิเคราะห์การถดถอบแบบสุ่ม (random - coefficients regression model) ตลอดจนไม่มีกระบวนการทดสอบความสอดคล้องของโมเดลกับข้อมูลเชิงประจักษ์ในคราวเดียวกัน ดังเช่นการวิเคราะห์ด้วยโปรแกรมลิสเรล (LISREL) หรือโปรแกรมที่ใช้วิเคราะห์ความสัมพันธ์โครงสร้างเชิงเส้นอื่น ๆ เช่น EQS, BMDP, EzPATH, LINC, LISCOMP, CALIS, COSAN และ AMOS เป็นต้น ดังนั้นเราจะสามารถทำการทดสอบสมมติฐานศูนย์ (null hypothesis, H_0) ที่ว่าโมเดลตามสมมติฐานสอดคล้องกับข้อมูลเชิงประจักษ์ โดยการใช้ (χ^2 test ระหว่างเมทริกซ์สหสัมพันธ์ที่คำนวณได้จากโมเดลกับเมทริกซ์สหสัมพันธ์ที่คำนวณได้จากข้อมูลเชิงประจักษ์ (observed correlation matrix) ที่องศาของความเป็นอิสระ เท่ากับจำนวนเส้นทางความสัมพันธ์เชิงสาเหตุที่สมมุติให้มีผลเป็นศูนย์ เมื่อ $\chi^2 = 0$ แสดงว่าเมทริกซ์สหสัมพันธ์ทั้งสองเมทริกซ์มีค่าเท่ากัน หรือโมเดลมีความสอดคล้องกับข้อมูลอย่างสมบูรณ์ แต่ถ้าปฏิเสธ H_0 แสดงว่าโมเดลไม่สอดคล้องกับข้อมูลเชิงประจักษ์ โดยผู้วิจัยได้ประยุกต์ขั้นตอนการทดสอบความสอดคล้องของโมเดลตามสมมุติฐานกับข้อมูลเชิงประจักษ์ ด้วยวิธีของสเปคมาใช้ในเทคนิควิธีที่พัฒนาขึ้น ดังนี้

3.1) หาค่า R_m^2 จากสูตร

$$R_m^2 = 1 - (1 - R_1^2)(1 - R_2^2) \dots (1 - R_p^2)$$

เมื่อ R_m^2 แทน Ordinary squared multiple correlation coefficient
ของโมเดลเชิงสาเหตุแบบเต็มรูป

3.2) หาค่า M จากสูตร

$$M = 1 - (1 - R_1^2)(1 - R_2^2) \dots (1 - R_p^2)$$

เมื่อ M แทน Ordinary squared multiple correlation coefficient ของโมเดลเชิงสาเหตุตามสมมติฐาน

โดย R_1^2 ถึง R_p^2 แทนค่าประสิทธิภาพการพยากรณ์ (Coefficient of determination) ที่คำนวณได้ จากผลการวิเคราะห์ด้วยโปรแกรมเอชแอลเอ็ม ดังนี้ (ศิริชัย กาญจนวาสี, 2535)

$$R^2 = \frac{\sigma_1^2 - \sigma_2^2}{\sigma_1^2}$$

เมื่อ R^2 คือ ประสิทธิภาพการพยากรณ์

σ_1^2 คือ within - unit variance จากการวิเคราะห์ขั้น Null Model

σ_2^2 คือ within - unit variance จากการวิเคราะห์ขั้น Simple Model

3.3) หาค่า Q ซึ่งเป็นค่าสถิติที่ใช้วัดความสอดคล้อง (measurement of goodness of fit) ของโมเดลเชิงสาเหตุตามสมมติฐานกับข้อมูลเชิงประจักษ์ จากสูตร

$$Q = \frac{1 - R^2_m}{1 - M}$$

3.4) หาค่า W ซึ่งเป็นค่าสถิติที่ใช้ทดสอบนัยสำคัญของค่า Q จากสูตร

$$W = -(N - d) \log_e Q$$

เมื่อ W แทนค่าสถิติทดสอบที่มีการแจกแจงเป็น χ^2 ซึ่งมี $df = d$

N แทนขนาดของกลุ่มตัวอย่าง

d แทนจำนวนเส้นทางที่ถูกระบุว่ามีความเป็นศูนย์ (จึงไม่ได้ลากเส้นทางนั้นเข้าไว้ในโมเดล) ในโมเดลเชิงสาเหตุตามสมมติฐาน

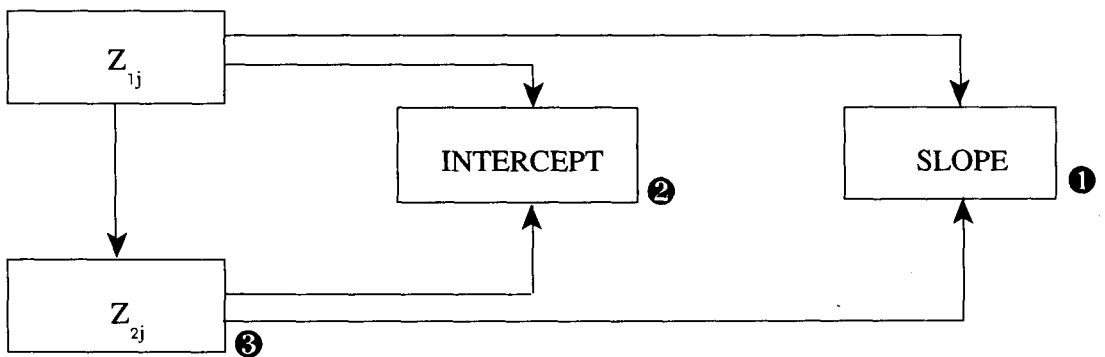
\log_e แทน natural logarithm

4.) ผลการพัฒนาวิธีวิเคราะห์แยกค่าสหสัมพันธ์ระหว่างตัวแปร ในโมเดลเชิงสาเหตุระดับนักเรียน ที่ผ่านการทดสอบความสอดคล้องกับข้อมูลเชิงประจักษ์แล้ว พบว่า การพิจารณาความสอดคล้องของโมเดลเชิงสาเหตุตามสมมติฐานกับข้อมูลเชิงประจักษ์นั้น พิจารณาจากการทดสอบนัยสำคัญของค่า Q จากค่า W ถ้าค่า W ไม่มีนัยสำคัญทางสถิติ หมายความว่าโมเดลเชิงสาเหตุตามสมมติฐานอธิบายระบบของความสัมพันธ์ได้ไม่แตกต่างจากโมเดลเชิงสาเหตุเต็มรูป ซึ่งแสดงว่าโมเดลเชิงสาเหตุตามสมมติฐาน มีความสอดคล้องกับข้อมูลเชิงประจักษ์ แต่ถ้า W มีนัยสำคัญทางสถิติ หมายความว่าโมเดลเชิงสาเหตุตามสมมติฐาน อธิบายความแปรปรวนระบบของ

ความสัมพันธ์ได้แตกต่างจากโมเดลเชิงสาเหตุเต็มรูป ซึ่งแสดงว่าโมเดลเชิงสาเหตุตามสมมติฐานยังไม่สอดคล้องกับข้อมูลเชิงประจักษ์อย่างมีนัยสำคัญ

ผลจากการวิเคราะห์โมเดลเชิงสาเหตุระดับนักเรียน หากพบว่า ค่าคงที่ (intercept) ของตัวแปรตาม และสัมประสิทธิ์การถดถอย (slope) จากการวิเคราะห์ถดถอยตามลำดับขั้นตอนด้วยโปรแกรมเอชแอลเอ็ม ค่าใดที่อิทธิพลสูงมีนัยสำคัญทางสถิติ ก็สมควรนำค่าดังกล่าว ไปเป็นตัวแปรตามในการวิเคราะห์ระดับชั้นเรียนต่อไป โดยมีมุ่งพิจารณาค่าคงที่ของตัวแปรตามที่อยู่ภายในสุดของระบบความสัมพันธ์ระดับนักเรียน และสัมประสิทธิ์การถดถอย ที่มีต่อตัวแปรตามดังกล่าวเป็นสำคัญ

1.2.2 ผลพัฒนาการวิเคราะห์เชิงสาเหตุแบบพหุระดับ ในระดับชั้นเรียน (causal macro model) พบว่า การวิเคราะห์ในขั้นตอนนี้ สามารถประยุกต์วิธีการวิเคราะห์ถดถอยเป็นลำดับขั้นตามโมเดลเชิงสาเหตุกับเทคนิคการวิเคราะห์พหุระดับด้วยโปรแกรมเอชแอลเอ็ม ในขั้นตอนของ Hypothetical Model โดยจากโมเดลการวิเคราะห์ที่นำเสนอไปแล้วในเบื้องต้น เมื่อทำการวิเคราะห์ต่อในระดับชั้นเรียนสามารถเขียนโมเดลการวิเคราะห์ได้ดังนี้



- ❶ วิเคราะห์ด้วยโปรแกรมเอชแอลเอ็ม ตัวแปรตามคือสัมประสิทธิ์การถดถอย (slope) ที่อิทธิพลสูงมีนัยสำคัญทางสถิติ จากการวิเคราะห์โมเดลเชิงสาเหตุระดับนักเรียน
- ❷ วิเคราะห์ด้วยโปรแกรมเอชแอลเอ็ม ตัวแปรตามคือค่าคงที่ (intercept) ที่อิทธิพลสูงมีนัยสำคัญทางสถิติ จากการวิเคราะห์โมเดลเชิงสาเหตุระดับนักเรียน
- ❸ วิเคราะห์ถดถอยด้วยโปรแกรม SPSS/PC * ตัวแปรตามคือ Z_{2j}

ภาพที่ 3 แสดง Causal Macro Model

จากภาพที่ 3 แสดงให้เห็นว่า การวิเคราะห์เชิงสาเหตุแบบพหุระดับ ในระดับชั้นเรียนนี้ จะนำค่าคงที่ (intercept) และค่าสัมประสิทธิ์เส้นทางความสัมพันธ์เชิงสาเหตุ หรือสัมประสิทธิ์การถดถอย (ซึ่งผลการวิเคราะห์หรืออิทธิพลสุ่มหรือ Random Effect จะต้องมีความผันแปรระหว่างหน่วยการวิเคราะห์อย่างมีนัยสำคัญทางสถิติ) มาเป็นตัวแปรตามโดยใช้ตัวแปรอิสระระดับในระดับชั้นเรียน คือ Z_{1j} และ Z_{2j} เป็นตัวพยากรณ์ ทั้งนี้การวิเคราะห์หรืออิทธิพลของตัวแปรระดับชั้นเรียนที่มีต่อค่าคงที่ (intercept) และค่าสัมประสิทธิ์การถดถอย ที่นำมาจากวิเคราะห์ระดับนักเรียนนั้น จะอยู่ในขั้นตอนการวิเคราะห์ Hypothetical Model ของโปรแกรมเอชแอลเอ็ม แต่การพิจารณาอิทธิพลของ Z_{1j} ต่อ Z_{2j} นั้น โปรแกรมเอชแอลเอ็มไม่สามารถวิเคราะห์ได้ ดังนั้นจึงต้องใช้โปรแกรมคอมพิวเตอร์มาตรฐานเช่น SPSS/PC วิเคราะห์ประกอบด้วยซึ่งจะสามารถวิเคราะห์ได้เฉพาะอิทธิพลคงที่ และเมื่อพิจารณาตามโมเดลการวิเคราะห์ที่นำเสนอแล้ว หากพบว่าอิทธิพลคงที่ (fixed effect) หรือเส้นทางความสัมพันธ์เชิงสาเหตุจาก Z_{1j} ไปยัง Z_{2j} มีนัยสำคัญทางสถิติ จะต้องพิจารณาในลักษณะของความสัมพัทธ์เชิงสาเหตุจาก Z_{1j} ที่ส่งผลต่อ Z_{2j} ก่อนส่งผ่านไปยังค่าคงที่ (intercept) และสัมประสิทธิ์การถดถอย (slope) ด้วย แต่หากพบว่าอิทธิพลคงที่ (fixed effect) หรือเส้นทางความสัมพันธ์เชิงสาเหตุจาก Z_{1j} ไปยัง Z_{2j} ไม่มีนัยสำคัญทางสถิติ ก็ไม่จำเป็นต้องพิจารณาเส้นทางความสัมพันธ์เชิงสาเหตุจาก Z_{1j} ที่ส่งผลต่อ Z_{2j} ก่อนส่งผ่านไปยังค่าคงที่ (intercept) และสัมประสิทธิ์การถดถอย (slope) แต่อย่างใด ดังนั้น จะสามารถนำเสนอผลการพัฒนาการวิเคราะห์ความสัมพันธ์เชิงสาเหตุแบบพหุระดับในระดับชั้นเรียนได้ดังนี้

1) ผลการพัฒนากการวิเคราะห์อิทธิพลคงที่ พบว่า ในส่วนของการวิเคราะห์ด้วยโปรแกรมเอชแอลเอ็ม ใช้หลักการเดียวกันกับการวิเคราะห์ในระดับนักเรียน โดยพิจารณาความมีนัยสำคัญทางสถิติของอิทธิพลคงที่ จากตัวแปร Z_{1j} และ Z_{2j} ในโมเดลความสัมพันธ์เชิงสาเหตุระดับชั้นเรียนที่ส่งผลต่อค่าคงที่ (intercept) และค่าสัมประสิทธิ์การถดถอย (slope) ซึ่งถูกยกกระดับขึ้นเป็นตัวแปรตามระดับชั้นเรียน ส่วนการวิเคราะห์ด้วยโปรแกรม SPSS/P* ยกตัวอย่างเช่นในโมเดลที่นำเสนอตามภาพที่ 3 นั้น หากอิทธิพลคงที่ของตัวแปร Z_{1j} ที่ส่งผลต่อ Z_{2j} มีนัยสำคัญทางสถิติ ก็แสดงว่าค่าคงที่ (intercept) และค่าสัมประสิทธิ์การถดถอย (slope) ได้รับอิทธิพลจาก Z_{1j} ที่ส่งผลผ่าน Z_{2j} ด้วย ทั้งนี้หากโมเดลเชิงสาเหตุระดับชั้นเรียนมีความซับซ้อนมากกว่าตัวอย่างที่นำเสนอ การวิเคราะห์อิทธิพลคงที่ก็ย่อมมีความซับซ้อนมากขึ้นตามไปด้วย

2) ผลการพัฒนากการวิเคราะห์อิทธิพลสุ่ม พบว่า เมื่อพิจารณาตามโมเดลเชิงสาเหตุระดับชั้นเรียนที่นำเสนอต่อภาพที่ 3 ในส่วนของการวิเคราะห์ด้วยโปรแกรมเอชแอลเอ็ม การวิเคราะห์อิทธิพลสุ่มจะเป็นการพิจารณาความผันแปรของค่าคงที่ (intercept) และค่าสัมประสิทธิ์การถดถอย (slope) ซึ่งเป็นตัวแปรตามภายในสุดของการวิเคราะห์ของเส้นทางความ

สัมพันธเชิงสาเหตุระดับชั้นเรียนเท่านั้น โดยจะไม่สามารถพิจารณาอิทธิพลร่วมของสัมประสิทธิ์ถดถอยของ Z_{1j} และ Z_{2j} ที่มีต่อค่าคงที่ (intercept) และค่าสัมประสิทธิ์การถดถอย (slope) ส่วนผลจากการวิเคราะห์อิทธิพลจาก Z_{1j} ที่ส่งผลต่อ Z_{2j} ด้วยโปรแกรม SPSS/PC⁺ ก็ไม่สามารถพิจารณาอิทธิพลร่วมได้แต่ประการใด

3) ผลการพัฒนาการทดสอบความสอดคล้องของโมเดลกับข้อมูลเชิงประจักษ์ การวิเคราะห์ในขั้นตอนนี้ อาศัยหลักการและวิธีการคำนวณเช่นเดียวกับการวิเคราะห์ระดับนักเรียน แต่มีวัตถุประสงค์แตกต่างกันคือ เป็นการวิเคราะห์เพื่อหาโมเดลระดับชั้นเรียนที่มีความสอดคล้องกับข้อมูลเชิงประจักษ์โดยมีค่าคงที่ (intercept) และสัมประสิทธิ์การถดถอยที่ได้จากการวิเคราะห์ในระดับนักเรียน เป็นตัวแปรตามเป็นประการสำคัญ ตลอดจนการคำนวณหาค่า R^2 โดยการวิเคราะห์ระดับชั้นเรียนนี้ จะใช้ทั้งค่า R^2 ที่ได้จากผลการวิเคราะห์ด้วย SPSS/PC⁺ และค่า R^2 ที่คำนวณจากผลการวิเคราะห์ด้วยเอชแอลเอ็ม ดังนี้ (ศิริชัย กาญจนวาสี, 2535)

$$R^2 = \frac{T_1 - T_2}{T_1}$$

เมื่อ R^2 คือ ประสิทธิภาพในการพยากรณ์

T_1 คือ parameter variance จากการวิเคราะห์ขั้น Simple Model

T_2 คือ parameter variance จากการวิเคราะห์ขั้น Hypothetical Model

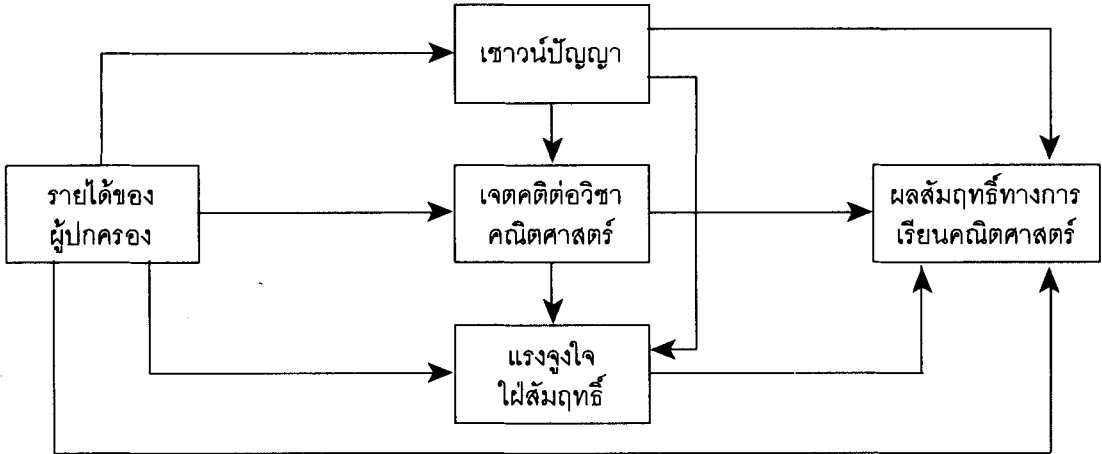
4) ผลการพัฒนาการวิเคราะห์แยกค่าสหสัมพันธ์ระหว่างตัวแปร เป็นการวิเคราะห์เพื่อพิจารณาผลกระทบทางตรง ผลกระทบทางอ้อมและผลกระทบรวมภายในโมเดลความสัมพันธเชิงสาเหตุระดับชั้นเรียน โดยอาศัยหลักการและวิธีการเช่นเดียวกับการวิเคราะห์แยกค่าสหสัมพันธ์ระหว่างตัวแปรของโมเดลเชิงสาเหตุระดับนักเรียนแทบทุกประการแต่สิ่งที่แตกต่างกันคือ การวิเคราะห์แยกค่าสหสัมพันธ์ของโมเดลเชิงสาเหตุระดับชั้นเรียนจะใช้ผลการวิเคราะห์จากโปรแกรม SPSS/PC⁺ ประกอบด้วย

ตอนที่ 2 ผลการทดลองใช้เทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับ

1) **ผลการวิเคราะห์ Null Model** เพื่อพิจารณาความผันแปรของตัวแปรตามแต่ละตัว ในโมเดลเชิงสาเหตุระดับนักเรียน ด้วยโปรแกรม เอชแอลเอ็ม

โมเดลเชิงสาเหตุระดับนักเรียนที่นำมาวิเคราะห์ ประกอบด้วยตัวแปรระดับนักเรียนที่ส่งผลต่อตัวแปรตาม ทั้งในแบบเต็มรูปและตามสมมติฐาน เมื่อประยุกต์โปรแกรม

เอชแอลเอ็มเข้ามาวิเคราะห์ จึงต้องวิเคราะห์ขั้น Null Model ก่อนเป็นขั้นตอนแรก โดยพิจารณาจากโมเดลเชิงสาเหตุแบบเต็มรูป ดังนี้



ภาพที่ 4 โมเดลเชิงสาเหตุระดับนักเรียนแบบเต็มรูป (Full Model)

การวิเคราะห์ Null Model สำหรับโมเดลเชิงสาเหตุระดับนักเรียนแบบเต็มรูป โดยใช้ผลสัมฤทธิ์ทางการเรียนเป็นตัวแปรตามในการวิเคราะห์ครั้งที่ 1 ใช้แรงจูงใจใฝ่สัมฤทธิ์ เจตคติต่อวิชาคณิตศาสตร์ และเซาว์นปัญญาเป็นตัวแปรตามในการวิเคราะห์ครั้งที่ 2, 3 และ 4 ตามลำดับ ทั้งนี้เพื่อพิจารณาความผันแปรของตัวแปรตามดังกล่าวโดยไม่มีตัวแปรอิสระใด ๆ ในระดับนักเรียนเข้ามาร่วมวิเคราะห์ ซึ่งจากการพิจารณาอิทธิพลสัมพัทธ์ ตัวแปรตามแต่ละตัวในระบบความสัมพันธ์มีความแปรผันระหว่างห้องเรียนอย่างมีนัยสำคัญทางสถิติ จึงสามารถนำไปวิเคราะห์ในขั้นตอนต่อไปได้

2) ผลการวิเคราะห์ Simple Model ผู้เขียนได้นำเสนอเฉพาะโมเดลตามสมมติฐานดังนี้

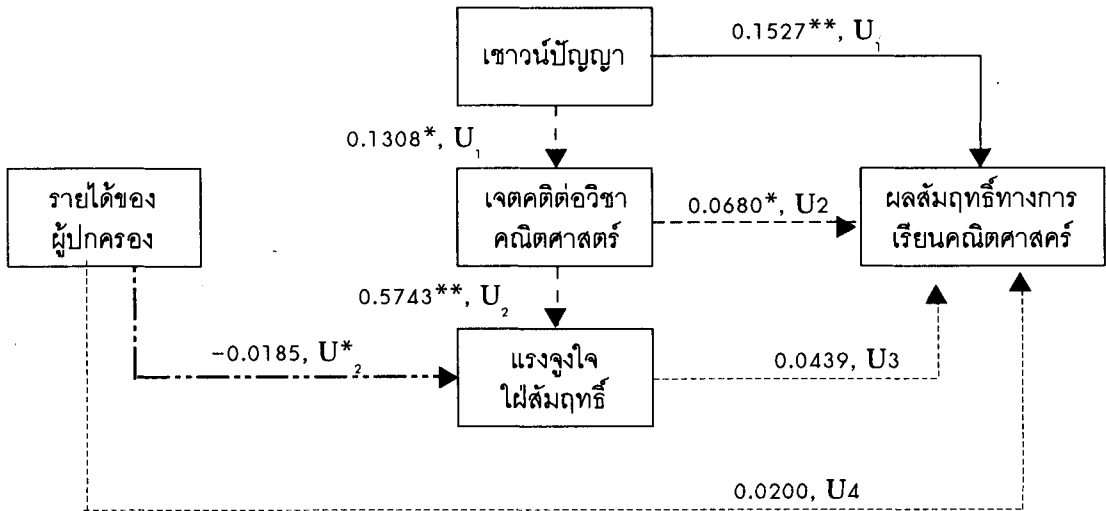
ตารางที่ 1 แสดงค่าสัมประสิทธิ์เส้นทางความสัมพันธ์เชิงสาเหตุและค่าสถิติที่ได้จากการวิเคราะห์โมเดลเชิงสาเหตุระดับนักเรียนตามสมมติฐาน

ตัวแปรสาเหตุ	ตัวแปรผล	R ²	อิทธิพลคงที่		อิทธิพลสุ่ม	
			P _{jk} (β)	t	U	χ ²
IQ	ACH	0.1005	0.1527**	4.038	U ₁ *	32.0814
ATI			0.0680*	2.202	U ₂	19.9270
MOTV			0.0439	1.389	U ₃	24.5143
INCOMP			0.0200	0.703	U ₄	8.5894
ATI	MOTV	0.3248	0.5734**	14.845	U ₁	19.3496
INCOMP			-0.0185	-0.354	U ₂ *	35.2316
IQ	ATI	0.0248	0.1308*	2.202	U ₁	28.9228

*p < 0.05, **p < 0.01

จากตาราง เมื่อนำค่าสัมประสิทธิ์ถดถอยมาตรฐาน หรือสัมประสิทธิ์เส้นทางความสัมพันธ์เชิงสาเหตุระหว่างตัวแปรในโมเดลเชิงสาเหตุระดับนักเรียนตามสมมติฐาน มาแสดงในลักษณะของแผนภาพ จะได้ภาพแสดงอิทธิพลของตัวแปรต่าง ๆ ในโมเดลเชิงสาเหตุระดับนักเรียนตามสมมติฐานทั้งในส่วนของอิทธิพลคงที่ และอิทธิพลสุ่ม ดังนี้

ภาพที่ 5 โมเดลเชิงสาเหตุระดับนักเรียน แสดงอิทธิพลของตัวแปรแรงจูงใจใฝ่สัมฤทธิ์ เจตคติ
 วิชาคณิตศาสตร์ เซาว์นปัญญา และรายได้ของผู้ปกครอง ที่มีต่อผลสัมฤทธิ์ทางการเรียน
 คณิตศาสตร์, ความมีนัยสำคัญทางสถิติของอิทธิพลคงที่และอิทธิพลสุ่มของสัมประสิทธิ์
 เส้นทางในระบบ ความสัมพันธ์ตามสมมติฐาน



- ▶ สัมประสิทธิ์เส้นทางที่อิทธิพลคงที่ และ อิทธิพลสุ่มไม่มีนัยสำคัญ
- ▶ สัมประสิทธิ์เส้นทางที่อิทธิพลคงที่ไม่มีนัยสำคัญ แต่ อิทธิพลสุ่มมีนัยสำคัญ
- ▶ สัมประสิทธิ์เส้นทางที่อิทธิพลคงที่มีนัยสำคัญ แต่ อิทธิพลสุ่มไม่มีนัยสำคัญ
- ▶ สัมประสิทธิ์เส้นทางที่อิทธิพลคงที่ และ อิทธิพลสุ่ม มีนัยสำคัญ

ค่าตัวเลข หมายถึงสัมประสิทธิ์เส้นทาง พิจารณาความมีนัยสำคัญของอิทธิพลคงที่ได้จาก เครื่องหมาย * หรือ ** (* p < 0.05 , ** p < 0.01)

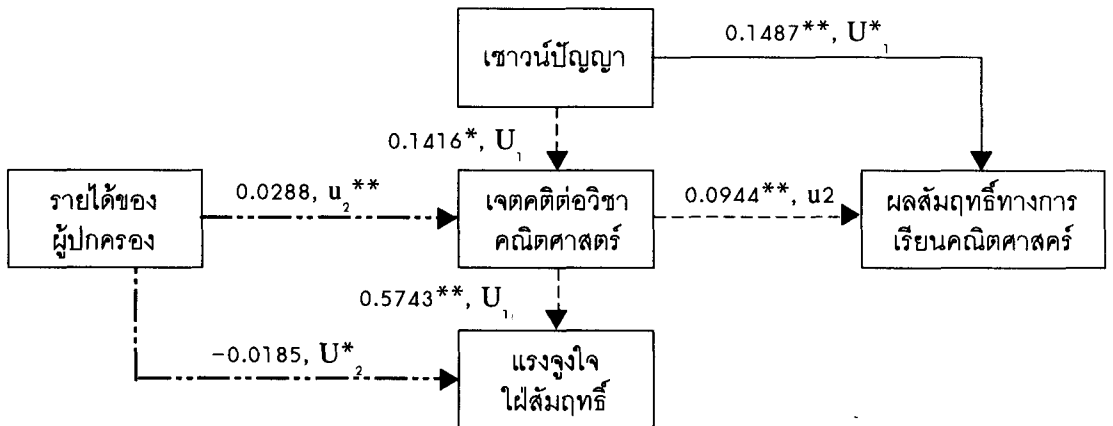
U หมายถึง อิทธิพลสุ่มของสัมประสิทธิ์เส้นทาง พิจารณาความมีนัยสำคัญได้จาก เครื่องหมาย * หรือ ** (* p < 0.05 , ** p < 0.01)

3) ผลการทดสอบความสอดคล้องของโมเดลเชิงสาเหตุระดับนักเรียนตามสมมติฐานกับข้อมูลเชิงประจักษ์ด้วยวิธีของสเปค

ผลการวิเคราะห์ พบว่าควรตัดเส้นทางที่ทั้งอิทธิพลคงที่และอิทธิพลสุ่มไม่มีนัยสำคัญทางสถิติออกจากโมเดล และควรเพิ่มเส้นทางที่อิทธิพลสุ่มมีนัยสำคัญทางสถิติ (ถึงแม้อิทธิพลคงที่จะ

ไม่มีนัยสำคัญทางสถิติ) เข้าในโมเดล เนื่องจากผลการพัฒนาเทคนิควิธีวิเคราะห์ พบว่า หากตัดเส้นทางดังกล่าวออกไป เพราะอิทธิพลคงที่ไม่มีนัยสำคัญทางสถิติ ตามหลักของการวิเคราะห์ด้วยเทคนิควิเคราะห์เชิงสาเหตุที่ใช้กันทั่วไปแล้ว จะทำให้ค่า W มีนัยสำคัญทางสถิติ นั่นคือ โมเดลการวิเคราะห์จะไม่สอดคล้องกับข้อมูลเชิงประจักษ์ และขาดความสมเหตุสมผล เพราะตัวแปรรายได้ของผู้ปกครอง จะถูกตัดออกจากโมเดลการวิจัย ดังนั้นจึงควรคงเส้นทางดังกล่าวไว้ เพื่อเป็นการยืนยันถึงความมั่นคงในทฤษฎีของผู้วิจัย ดังนั้น เมื่อทำการวิเคราะห์ใหม่จะได้โมเดลเชิงสาเหตุระดับนักเรียนที่ปรับใหม่ แสดงตามภาพที่ 6 ดังนี้

ภาพที่ 6 โมเดลเชิงสาเหตุระดับนักเรียน แสดงอิทธิพลของตัวแปรแรงจูงใจใฝ่สัมฤทธิ์ เจตคติต่อวิชาคณิตศาสตร์ เขาวนปัญญา และรายได้ของผู้ปกครอง ที่มีต่อผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์, ความมีนัยสำคัญทางสถิติของอิทธิพลคงที่และอิทธิพลสุ่มของสัมประสิทธิ์เส้นทางในระบบความสัมพันธ์ที่ปรับปรุงใหม่



อนึ่ง โมเดลเชิงสาเหตุระดับนักเรียนที่ปรับปรุงใหม่นี้ เป็นโมเดลที่มีความเหมาะสมสำหรับใช้วิเคราะห์แยกค่าสหสัมพันธ์ระหว่างตัวแปร เพื่อพิจารณาผลกระทบทางตรง ผลกระทบทางอ้อม และผลกระทบรวมของตัวแปรอิสระในโมเดลที่มีต่อตัวแปรตาม ตลอดจนใช้วิเคราะห์ในระดับชั้นเรียน (macro level analysis) ต่อไป

4) ผลการวิเคราะห์อิทธิพลคงที่ อิทธิพลสุ่ม และการแยกค่าสหสัมพันธ์ระหว่างตัวแปรในโมเดลเชิงสาเหตุระดับนักเรียน ที่สอดคล้องกับข้อมูลเชิงประจักษ์

ผลการวิเคราะห์ในขั้นตอนนี้ ผู้เขียนได้นำเสนอเฉพาะ โมเดลที่ใช้ผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์เป็นตัวแปรตาม ดังนี้

ตารางที่ 2 แสดงอิทธิพลคงที่ อิทธิพลสุ่ม ของการวิเคราะห์อิทธิพลภายในห้องเรียน (pooled within class effect) และความแปรปรวนระหว่างห้องเรียน (between class variance) เมื่อนำตัวแปรชาวน์ปัญญา (IQ), เจตคติต่อวิชาคณิตศาสตร์ (ATI) มาวิเคราะห์ร่วม ในสมการ โดยมีผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ (ACH) เป็นตัวแปรตาม (ค่าในวงเล็บคือสถิติทดสอบ t-test)

ตัวแปร	Fixed Effect		Random Effect			
	Pooled - Within Class Effect		Between - Class Variance		χ^2	df
ระดับนักเรียน	γ	β	Parameter Variance	Total Observed Variance		
intercept (γ_{00})	12.5084** (12.149)		21.8272	34.2528	1455.5995**	20
IQ	0.1343** (3.983)	0.1487**	0.0098	12.4354	35.6635*	20
ATI	0.0318** (3.535)	0.0944**	0.0003	12.4259	25.4984	20
within unit error variance		12.4256				
$R^2 = 0.0878$						

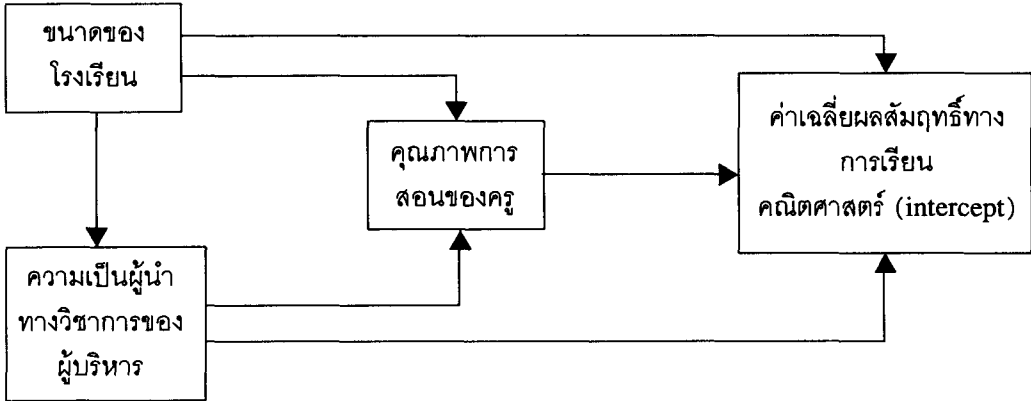
*P < 0.05, **P < 0.01

จากตารางที่ 2 การพิจารณาอิทธิพลคงที่ (fixed effect) พบว่า ค่าคงที่ของการวิเคราะห์ (γ_{00}) สัมประสิทธิ์การถดถอยของชาวน์ปัญญา (γ_{IQ}) และสัมประสิทธิ์การถดถอยของตัวแปรเจตคติต่อวิชาคณิตศาสตร์ (γ_{ATI}) มีนัยสำคัญทางสถิติที่ระดับ 0.01 ต่อผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ของนักเรียน เมื่อพิจารณาอิทธิพลสุ่ม (random effect) พบว่า ค่าคงที่ (γ_{00}) มีความผันแปรระหว่างห้องเรียนอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 ($\chi^2 = 1455.5995$) ส่วนสัมประสิทธิ์การถดถอยของตัวแปรชาวน์ปัญญา (γ_{IQ}) มีความผันแปรระหว่างห้องเรียนอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 ($\chi^2 = 35.6635$) ส่วนสัมประสิทธิ์การถดถอยของตัวแปรเจตคติต่อวิชาคณิตศาสตร์ (γ_{ATI}) ไม่มีความผันแปรระหว่างห้องเรียนอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05

สำหรับการวิเคราะห์แยกค่าสหสัมพันธ์ระหว่างตัวแปรในโมเดลเชิงสาเหตุระดับนักเรียนที่ปรับปรุงใหม่นั้น โดยปกติแล้วจะนำค่าสัมประสิทธิ์ถดถอยมาตรฐานหรือค่าสัมประสิทธิ์เส้นทางของตัวแปรอิสระที่มีนัยสำคัญทางสถิติ มาพิจารณาอิทธิพลที่มีต่อตัวแปรตาม ทั้งนี้ผลการวิเคราะห์จากโปรแกรมคอมพิวเตอร์มาตรฐานทั่ว ๆ ไป นั้น จะพิจารณาความมีนัยสำคัญทางสถิติเฉพาะอิทธิพลคงที่ หากสัมประสิทธิ์เส้นทางไม่มีนัยสำคัญทางสถิติ ก็จะตัดเส้นทางนั้นออกจากโมเดลการวิเคราะห์ แต่เมื่อใช้เทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับด้วยโปรแกรมเอชแอลเอ็มที่ผู้วิจัยพัฒนาขึ้นมานั้น ค่าอิทธิพลจากสัมประสิทธิ์เส้นทางที่ได้ จะประกอบไปด้วยทั้งอิทธิพลคงที่ (fixed effect) และอิทธิพลสุ่ม (random effect) ซึ่งผลรวมกันอยู่ตามหลักการของโมเดลสัมประสิทธิ์ถดถอยแบบสุ่ม (random-coefficient regression model) (Raudenbush and Bryk , 1992) ดังนั้นการพิจารณาความมีนัยสำคัญของสัมประสิทธิ์เส้นทาง จึงต้องพิจารณาอิทธิพลสุ่มประกอบกันไปด้วย โดยจะนำสัมประสิทธิ์เส้นทางที่อิทธิพลคงที่มีนัยสำคัญทางสถิติไปวิเคราะห์แยกค่าสหสัมพันธ์ระหว่างตัวแปรเป็นสำคัญ หากอิทธิพลคงที่ไม่มีนัยสำคัญทางสถิติ แต่อิทธิพลสุ่มมีนัยสำคัญทางสถิติ ก็สมควรนำไปวิเคราะห์แยกค่าสหสัมพันธ์ระหว่างตัวแปรในโมเดลได้ เนื่องจากผลการทดสอบความสอดคล้องของโมเดลกับข้อมูลเชิงประจักษ์ ซึ่งใช้ค่า R^2 จากโมเดลที่ยังคงเส้นทางดังกล่าวไว้ พบว่า มีความสอดคล้องกับข้อมูลเชิงประจักษ์ มากกว่าโมเดลที่ตัดเส้นทางดังกล่าวออกไป

๕) ผลการวิเคราะห์โมเดลเชิงสาเหตุระดับชั้นเรียน เพื่อพิจารณาอิทธิพลของตัวแปรระดับโรงเรียน ที่มีต่อค่าคงที่ (intercept) และสัมประสิทธิ์การถดถอย (slope)

จากหลักการของการวิเคราะห์ถดถอยแบบสุ่ม (random-coefficient regression model) การวิเคราะห์โมเดลเชิงสาเหตุระดับชั้นเรียน (causal macro model) จึงสามารถที่จะประมาณค่าความแปรปรวนของค่าเฉลี่ยของตัวแปรตามระดับนักเรียน (intercept) และค่าความแปรปรวนของสัมประสิทธิ์ถดถอย (slope) ด้วยโมเดลที่ใช้ค่าเฉลี่ยและความชันเป็นตัวแปรตาม (intercepts and slopes as outcome model) (Raudenbush and Bryk , 1992) ดังนี้



ภาพที่ 7 โมเดลเชิงสาเหตุระดับชั้นเรียนแบบเต็มรูป (full model) เมื่อใช้ค่าคงที่ของผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ (ACH-intercept) เป็นตัวแปรตาม

จากภาพที่ 7 แสดงว่า การวิเคราะห์โมเดลเชิงสาเหตุระดับชั้นเรียนนี้ ค่าคงที่ หรือค่าเฉลี่ยผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ จะถูกยกระดับขึ้นมาเป็นตัวแปรตามระดับชั้นเรียน หากทำการวิเคราะห์อิทธิพลเชิงสาเหตุของตัวแปรอิสระระดับชั้นเรียน ที่มีต่อตัวแปรตามดังกล่าว ด้วยขั้นตอนของ Hypothetical Model ของโปรแกรมเอชแอลเอ็มแล้ว กระบวนการวิเคราะห์ของโปรแกรมจะสามารถพิจารณาเฉพาะอิทธิพลเชิงสาเหตุ จากตัวแปรระดับโรงเรียนไปยังค่าเฉลี่ยผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ของนักเรียน เท่านั้น จะไม่สามารถวิเคราะห์อิทธิพลเชิงสาเหตุระหว่างตัวแปรอิสระดังกล่าวได้ ดังนั้น จึงต้องใช้โปรแกรม SPSS / PC* ประกอบการวิเคราะห์ด้วย ผลการวิเคราะห์ปรากฏดังตารางที่ 3 ดังนี้

ตารางที่ 3 ประมาณค่าอิทธิพลของตัวแปรขนาดของโรงเรียน (SIZES), ความเป็นผู้นำทางวิชาการของผู้บริหาร (HEAD) และคุณภาพการสอนของครู (QUALT) ที่มีต่อค่าเฉลี่ยผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ของนักเรียน (ACH - intercept)

ตัวแปรระดับชั้นเรียน	ประมาณค่าอิทธิพล ที่มีต่อ ACH - intercept		
	γ	β	t
Fixed Effect			
intercept	12.4947**	-	20.7913
QUALT	1.7522*	0.6757*	2.508
SIZES	1.4755	0.4750	1.904
HEAD	0.2882**	0.7843**	3.031
Random Effect			
Parameter Variance	7.1359		
Total Observed Variance	19.9558		
$\chi^2 = 304.0558^{**}$, df = 17			
Within - unit Variance	12.8199		
$R^2 = 0.6731$			

*p < 0.05, **p < 0.01

จากตารางที่ 3 การพิจารณาอิทธิพลคงที่ พบว่าค่าคงที่ของการวิเคราะห์ระดับชั้นเรียนซึ่งเป็นค่าเฉลี่ยของค่าเฉลี่ยของผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ มีอิทธิพลต่อตัวแปรตามอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 (t = 20.791) ส่วนตัวแปรระดับชั้นเรียนที่มีอิทธิพลอย่างมีนัยสำคัญทางสถิติ ที่ระดับ 0.01 คือ ความเป็นผู้นำทางวิชาการของผู้บริหาร (t = 3.031) ตัวแปรที่มีอิทธิพลอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 คือ คุณภาพการสอนของครู (t = 2.508) ส่วนตัวแปรขนาดของโรงเรียน อิทธิพลไม่มีนัยสำคัญทางสถิติที่ระดับ 0.05 เมื่อพิจารณาอิทธิพลสุ่มพบว่า ค่าเฉลี่ยของผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ (ACH - intercept) ยังมีความผันแปรระหว่างห้องเรียนอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 ($\chi^2 = (304.0558)$) และเมื่อวิเคราะห์อิทธิพลของตัวแปรอิสระ ต่อตัวแปรตามในระบบความสัมพันธ์ให้ครบถ้วน จะสามารถแสดงค่าสถิติจากการวิเคราะห์ ได้ดังนี้

ตารางที่ 4 แสดงค่าสัมประสิทธิ์เส้นทางความสัมพันธ์เชิงสาเหตุและค่าสถิติที่ได้จากการวิเคราะห์โมเดลเชิงสาเหตุระดับชั้นเรียนเมื่อใช้ค่าเฉลี่ยของผลสัมฤทธิ์ทางการเรียนวิชาคณิตศาสตร์ (ACH – intercept) เป็นตัวแปรตาม แบบเต็มรูป

ตัวแปรระดับชั้นเรียน	ตัวแปรผล	R ²	P _{jk} (β)	t
QUALT	ACH-intercept	0.6731	0.6757*	2.508
SIZES			0.4750	1.904
HEAD			0.7843**	3.031
HEAD	QUALT	0.2945	0.4188**	12.038
SIZES			0.2383**	6.851
SIZES	HEAD	0.0975	0.3122**	8.359

*P < 0.05, **P < 0.01

ผลการวิเคราะห์โมเดลเชิงสาเหตุระดับชั้นเรียนข้างต้น ผู้เขียนนำเสนอเฉพาะโมเดลแบบเต็มรูปเท่านั้น ทั้งนี้ การวิเคราะห์โมเดลเชิงสาเหตุระดับชั้นเรียนตามสมมติฐาน จะอาศัยหลักการวิเคราะห์เช่นเดียวกันกับการวิเคราะห์โมเดลเชิงสาเหตุระดับชั้นเรียนแบบเต็มรูป ซึ่งจะสามารถนำผลการวิเคราะห์ที่ได้ ไปทดสอบความสอดคล้องของโมเดลกับข้อมูลเชิงประจักษ์ ด้วยวิธีของสเปคและวิเคราะห์แยกค่าสหสัมพันธ์ระหว่างตัวแปรได้

อนึ่งสำหรับการวิเคราะห์โมเดลเชิงสาเหตุระดับชั้นเรียนโดยใช้สัมประสิทธิ์การถดถอย (slope) เป็นตัวแปรตามนั้น จะให้หลักการและกระบวนการวิเคราะห์เช่นเดียวกันกับโมเดลเชิงสาเหตุระดับชั้นเรียนที่ใช้ค่าเฉลี่ย หรือค่าคงที่ (intercept) ที่ผู้เขียนนำเสนอข้างต้นทุกประการ

ผลการวิเคราะห์โมเดลเชิงสาเหตุแบบพหุระดับที่นำเสนอไปข้างต้น มีประเด็นสำคัญที่สมควรพิจารณาประกอบด้วยทุกครั้ง คือ ในการพิจารณาตัดสินทางความสัมพันธ์เชิงสาเหตุที่ไม่มีนัยสำคัญทางสถิติออกจากโมเดลการวิเคราะห์ หรือการคงเส้นทางความสัมพันธ์เชิงสาเหตุที่มีนัยสำคัญทางสถิติไว้ในโมเดลการวิเคราะห์นั้น เนื่องจากเทคนิควิธีที่ผู้วิจัยพัฒนาขึ้น สามารถพิจารณาได้ทั้งอิทธิพลคงที่ของสัมประสิทธิ์เส้นทางตลอดจนอิทธิพลสุ่มหรือความผันแปรของสัมประสิทธิ์เส้นทางดังกล่าวควบคู่ไปพร้อม ๆ กัน ดังนั้น ผู้วิจัยจึงเสนอเกณฑ์การพิจารณา ดังนี้

1.) ในโมเดลเชิงสาเหตุระดับนักเรียน (causal micro model)

- 1.1) สัมประสิทธิ์เส้นทางที่อิทธิพลคงที่ และ อิทธิพลสุ่มไม่มีนัยสำคัญทางสถิติ ให้ตัดออกจากโมเดลการวิเคราะห์
- 1.2) สัมประสิทธิ์เส้นทางที่อิทธิพลคงที่ไม่มีนัยสำคัญแต่อิทธิพลสุ่มมีนัยสำคัญทางสถิติ ให้คงไว้ในโมเดลการวิเคราะห์
- 1.3) สัมประสิทธิ์เส้นทางที่อิทธิพลคงที่มีนัยสำคัญแต่อิทธิพลสุ่มไม่มีนัยสำคัญทางสถิติ ให้คงไว้ในโมเดลการวิเคราะห์
- 1.4) สัมประสิทธิ์เส้นทางที่อิทธิพลคงที่ และ อิทธิพลสุ่ม มีนัยสำคัญทางสถิติให้คงไว้ในโมเดลการวิเคราะห์

2.) ในโมเดลเชิงสาเหตุระดับชั้นเรียน (causal macro model)

นัยสำคัญทางสถิติของสัมประสิทธิ์เส้นทางที่ได้จากการวิเคราะห์ในโมเดลเชิงสาเหตุระดับชั้นเรียน จะสามารถพิจารณาได้เฉพาะอิทธิพลคงที่ ดังนั้นการพิจารณาตัดเส้นทางหรือคงเส้นทางไว้ในโมเดล จะสามารถพิจารณาได้จากนัยสำคัญทางสถิติของ การทดสอบ t เช่นเดียวกับการวิเคราะห์เชิงสาเหตุทั่ว ๆ ไป

อภิปรายผลการวิจัย

จากสรุปผลการวิจัยที่นำเสนอไปข้างต้น พบว่าเทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับ (multilevel causal analysis technique) ที่ผู้วิจัยพัฒนาขึ้นจากแนวคิดของ Jan de Leeuw (อ้างถึงใน Raudenbush and Bryk, 1992) โดยผู้วิจัยประยุกต์โปรแกรมเอชแอลเอ็มมาใช้วิเคราะห์นั้น มีกระบวนการวิเคราะห์และให้สารสนเทศที่น่าสนใจหลายประการ ดังนี้

1. การวิเคราะห์เชิงสาเหตุ (path analysis) สามารถประยุกต์ใช้ในการวิเคราะห์ข้อมูลแบบพหุระดับ (multilevel data) ได้โดยใช้เทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับ (multilevel causal analysis technique) ข้อค้นพบดังกล่าวนี้ สอดคล้องกับสมมติฐานข้อที่ 1 ของการวิจัย ทั้งนี้เนื่องมาจากโปรแกรมเอชแอลเอ็มที่ประยุกต์มาใช้วิเคราะห์นั้น แม้จะมีวัตถุประสงค์เพื่อวิเคราะห์ข้อมูลพหุระดับเพียงอย่างเดียว แต่ก็อาศัยหลักการวิเคราะห์ถดถอย (regression analysis) เป็นพื้นฐานสำคัญ ดังนั้นเมื่อนำมาประยุกต์ใช้วิเคราะห์เชิงสาเหตุ จึงสามารถวิเคราะห์ได้โดยการวิเคราะห์ถดถอยตัวแปรตามแต่ละตัวในโมเดลเชิงสาเหตุตามลำดับขั้นตอน เช่นเดียวกับการใช้เทคนิคการวิเคราะห์ถดถอยพหุคูณ (multiple regression) ด้วยโปรแกรม SPSS/PC* จึงอาจกล่าวได้ว่าเทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับนี้นอกจากจะพิจารณาถึงอิทธิพลเชิงสาเหตุระหว่างตัว

แปรในระบบความสัมพันธ์ได้แล้ว ยังสามารถพิจารณาโครงสร้างของข้อมูลที่แบ่งเป็นระดับตามสภาพธรรมชาติได้ด้วย

2. ผลการทดลองใช้เทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับ โดยการประยุกต์ใช้โปรแกรมเอชแอลเอ็ม สามารถพิจารณาได้กระบวนการวิเคราะห์ และสารสนเทศที่ได้ ดังนี้

2.1 เทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับ สามารถวิเคราะห์อิทธิพลคงที่ (fixed effects) อิทธิพลสุ่ม (random effects) หรือความแปรปรวนระหว่างห้องเรียน (between-class variance) โดยการพิจารณาความมีนัยสำคัญของการทดสอบ t และการทดสอบ χ^2 ตามรูปแบบของการวิเคราะห์ในโมเดลเอชแอลเอ็ม ซึ่งสอดคล้องกับสมมติฐานข้อที่ 2.1 ของการวิจัย

2.2 เทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับ สามารถทดสอบความสอดคล้องของโมเดลกับข้อมูลเชิงประจักษ์ได้ ทั้งในโมเดลการวิเคราะห์ระดับนักเรียน และระดับชั้นเรียน ข้อค้นพบดังกล่าวนี้ เป็นไปตามสมมติฐานข้อที่ 2.2 ของการวิจัย ทั้งนี้เนื่องจากผลการวิเคราะห์เชิงสาเหตุแบบพหุระดับทั้งในระดับนักเรียน (causal micro level) และระดับชั้นเรียน (causal macro level) สามารถคำนวณค่า R^2 สำหรับใช้ในการคำนวณตามสูตรของสเปค (Specht, 1975 อ้างถึงใน Pedhazur, 1982) เพื่อทดสอบความสอดคล้องของโมเดลเชิงสาเหตุทั้งสองระดับกับข้อมูลเชิงประจักษ์ได้

2.3 เทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับสามารถวิเคราะห์แยกค่าสหสัมพันธ์ระหว่างตัวแปร เพื่อศึกษาผลกระทบทางตรง (direct effect) ผลกระทบทางอ้อม (indirect effect) และผลกระทบรวม (total effects) ได้ทั้งการวิเคราะห์ระดับนักเรียน และระดับชั้นเรียน ข้อค้นพบดังกล่าวนี้ เป็นไปตามสมมติฐานข้อที่ 2.3 ของการวิจัย ทั้งนี้เนื่องจาก สามารถทดสอบความสอดคล้องของโมเดลกับข้อมูลเชิงประจักษ์ได้ ทั้งในโมเดลการวิเคราะห์ระดับนักเรียน และระดับชั้นเรียน จนได้โมเดลโมเดลสุดท้ายหรือโมเดลแต่งใหม่ (trimmed model) ที่มีลักษณะประหยัดและมีความเหมาะสมที่จะใช้วิเคราะห์แยกค่าสหสัมพันธ์ระหว่างตัวแปรในโมเดลเชิงสาเหตุแต่ละระดับได้

ข้อเสนอแนะในการนำเทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับไปใช้

1. เนื่องจากเทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับโดยการประยุกต์ใช้โปรแกรมเอชแอลเอ็มที่ผู้วิจัยพัฒนาขึ้นมา นั้น เป็นเทคนิคที่มีความซับซ้อนในการวิเคราะห์ ดังนั้นหากทำการวิเคราะห์ด้วยเทคนิควิธีดังกล่าว ผู้วิจัยจึงควรศึกษาให้เกิดความเข้าใจทั้งในเรื่องการวิเคราะห์พหุระดับ และการวิเคราะห์เชิงสาเหตุ ตลอดจนหลักการและขั้นตอนของการวิเคราะห์เชิงสาเหตุแบบพหุระดับ ทั้งนี้เพื่อให้การวิเคราะห์ข้อมูล และการแปลผลการวิเคราะห์ดำเนินไปอย่างถูกต้อง

2. ควรขยายขอบเขตของเทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับโดยการนำไปวิเคราะห์ข้อมูลพหุระดับทางมนุษยศาสตร์หรือสาขาอื่น ทั้งนี้เพื่อเป็นทางเลือกใหม่ของนักวิจัยในการวิเคราะห์ข้อมูล

3. การพิจารณาตัดสินเส้นทางหรือคงเส้นทางความสัมพันธ์เชิงสาเหตุ ในโมเดลเชิงสาเหตุแต่ละระดับ มีเกณฑ์การพิจารณาบางประการที่แตกต่างจากการวิเคราะห์เชิงสาเหตุทั่วไป ดังนั้นหากทำการวิเคราะห์ข้อมูลด้วยเทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับที่ผู้วิจัยพัฒนาขึ้น จึงควรยึดเกณฑ์การพิจารณาตามที่ผู้วิจัยนำเสนอ กล่าวคือ ในโมเดลเชิงสาเหตุระดับชั้นเรียน ควรตัดเฉพาะเส้นทางที่ทั้งอิทธิพลคงที่และอิทธิพลสุ่มไม่มีนัยสำคัญทางสถิติ ออกจากโมเดลเท่านั้น ส่วนโมเดลเชิงสาเหตุระดับชั้นเรียน สามารถใช้เกณฑ์การพิจารณาได้ดังเช่นเกณฑ์การพิจารณาของการวิเคราะห์เชิงสาเหตุที่นิยมวิเคราะห์กันทั่วไป

ข้อเสนอแนะในการทำวิจัยครั้งต่อไป

1. ควรทำการวิจัยโดยใช้เทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับ ด้วยการให้ข้อมูลที่เก็บรวบรวมขึ้นมาจริง มีกลุ่มตัวอย่างขนาดใหญ่ขึ้น และใช้ตัวแปรจำนวนมากขึ้นในโมเดลการวิเคราะห์แต่ละระดับ ทั้งนี้เพื่อให้ผลการวิจัยมีความถูกต้องสมบูรณ์มากขึ้น

2. ควรประยุกต์โปรแกรมคอมพิวเตอร์ที่มีประสิทธิภาพสูงในการวิเคราะห์เชิงสาเหตุ เช่น LISREL มาวิเคราะห์เชิงสาเหตุแบบพหุระดับหรือทำการวิเคราะห์เปรียบเทียบกับโปรแกรมอื่น ๆ เช่น VARCL หรือ ML/2 เป็นต้น

3. ควรทำการวิเคราะห์เชิงสาเหตุแบบพหุระดับ โดยการประยุกต์ใช้โปรแกรมเอชแอลเอ็มกับข้อมูลสามระดับ

4. ควรมีการพัฒนาโปรแกรมคอมพิวเตอร์สำหรับใช้วิเคราะห์เชิงสาเหตุแบบพหุระดับได้โดยตรง หรือพัฒนาโปรแกรมคอมพิวเตอร์ที่สามารถนำค่าสถิติจากเทคนิควิธีวิเคราะห์เชิงสาเหตุแบบพหุระดับที่ผู้วิจัยพัฒนาขึ้น ไปประมวลผลได้อย่างรวดเร็วและถูกต้อง

เอกสารอ้างอิง

- ประเสริฐ เตชะนาราเกียรติ .**ความสัมพันธ์ระหว่างองค์ประกอบด้านนักเรียน องค์ประกอบด้าน
ครูสภาพแวดล้อมทางบ้าน และสภาพแวดล้อมทางโรงเรียน กับผลสัมฤทธิ์ทางการเรียน
คณิตศาสตร์ของนักเรียนชั้นมัธยมศึกษาปีที่4กรุงเทพมหานคร.** วิทยานิพนธ์ปริญญา
มหาบัณฑิต บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย , 2532 .
- ศิริชัย กาญจนวาสิ . การวิเคราะห์พหุระดับสำหรับการวิจัยทางการศึกษา . **ข่าวสารวิจัยการศึกษา**
5 (มิถุนายน - กรกฎาคม 2535) : 6-14 .
- Aitkin, M., and Longford, N. Statistical Modelling Issues in School Effectiveness
Studies . **Journal of the Royal Statistical Society.** (Series A). 1986, 149,1-26.
- Ann M. Gallagher and Richard De Lisi 'Gender Differences on Scholastic Aptitude Test
Mathematics Problem Solving among High - Ability Students' **Journal of
Educational Psychology**, 86(2) , 1994.
- Burstein, L.Analysis of Multilevel Data in Educational Research and Evaluation. In
Bertiner (Ed.) **Review of Research in Education.** Washington, D.C. :American
Educational Research Association , 1980 a.
- Burstein, L. The Role of Level of Analysis in the Specification of Education Effects .
In R.Dreeben and J.A Thomas (Eds.). **The Analysis of Educational Product-
ivity : Microanalysis.** Massachusetta : Ballinger Publishing Company, 1980 b.
- Burstein, L ; Linn, R.L ; and Capell I . Applying Multilevel Data in the Presence of
Heterogeneous Within - Class Regressions. **Journal of Educational Statistics** .
1978, 4, 347-89.
- Cronbach , L.J. **Research on Classrooms and School : Formulation of Questions, Design
and Analysis.** Occasional Paper, Stanford Evaluation Consortium, July ,1976 .
- Kreft , I.G.G.; de Leeuw , J. Questioning Multilevel Models . **Journal of Educational and
Behavioral Statistics.** (1995) , 2 171-189.
- Kanjanawasee , S. **Alternative Strategies for Policy Analysis : An Assessment of School
Effects on Students's Cognitive and Affective Mathematics Outcomes in
Lower Secondary School in Thailand.** Doctoral Dissertation in Education ,
University of California, Los Angeles, 1989.

- Mason , W.M. ; Wong G.Y. ; and Entwistle, B. Contextual Analysis through the Multilevel Linear Model. In S.Leinhardt , (Ed.) *Sociological Methodology 1983 - 1984*. (pp. 72-103). San Francisco : Jossey - Bass , 1983.
- Pedhazur , E.J. *Multiple Regression in Behavioral Research* . New York : Holt, Rinehart , and Winston, 1982.
- Raudenbush , S.W., Chinapat and Mohamed, K. Predictors and Consequences of Primary Teacher's Sense of Efficacy and Students' Perceptions of Teaching Quality in Thailand. *International Journal of Educational Research*. 1992, 17(2) : 165-177.
- Raudenbush, S.W. and Bryk, A.S. A Hierarchical Linear Model for Studying School Effects, *Sociology of Education*. 1986, 59, 1-17.
- Raudenbush, S.W., and Bryk, A.S. A Hierarchical Linear Model : *Applications and Data Analysis Methods*. Newbury Park : Sage Publication, 1992.

The Evaluation of the Cooperation Project between The Government of Thailand and the Asian Institute of Technology in Producing Personnel in Engineering and Technological Fields

Boonreang Kajornsinn¹ Samnao Kajornsinn¹ Pornthip Chaiyaso¹
Panurut Rattayapas² Tasanee Sahawat² Wantani Thaitiang²

ABSTRACT

The main purposes of the evaluation of cooperation project between the government of Thailand and the Asian Institute of Technology in producing personnel in engineering and technological fields are as follows :

(1) to follow up and evaluate if the steps taken in producing personnel in engineering and technological fields have met the goals.

(2) to follow up the graduates' application of their knowledge in developing the country.

(3) to evaluate their involvement in tackling the problems of inadequate personnel in engineering and technological fields.

The results indicated that the cooperation project between the Government of Thailand and the Asian Institute of Technology in producing personnel in engineering and technological fields had been on target with more than fifty percent of those sharing the solutions to such problems.

¹ Faculty of Education, Kasetsart University

² Office of the National Education Commission

การประเมินผลความร่วมมือระหว่างรัฐบาลไทย และสถาบันเทคโนโลยีแห่งเอเชีย ในโครงการพัฒนาบุคลากรทางวิศวกรรมศาสตร์ และเทคโนโลยี

บุญเรียง ขจรศิลป์¹ สำเนาวิ ขจรศิลป์¹ พรทิพย์ ไชยโล¹
ภาณุรัตน์ รัตยาภาส² ทศนีย์ สหวัฒน์² วันทนีย์ ไทยเที่ยง²

บทคัดย่อ

การประเมินผลความร่วมมือระหว่างรัฐบาลไทยและสถาบันเทคโนโลยีแห่งเอเชีย
ในโครงการพัฒนาบุคลากรทางวิศวกรรมศาสตร์ และเทคโนโลยี มีวัตถุประสงค์เพื่อ
1) ติดตามและประเมินการดำเนินงานในการผลิตบุคลากรตามเป้าหมาย 2) ติดตาม
การนำความรู้ความสามารถของบัณฑิตที่สำเร็จการศึกษาไปใช้ในการพัฒนาประเทศ
3) ประเมินการมีส่วนร่วมในการแก้ปัญหาการขาดแคลนอัตรากำลังคนด้านวิศวกรรม
ศาสตร์และเทคโนโลยี

ผลการประเมินปรากฏว่า โครงการฯ สามารถผลิตบัณฑิตได้ตามเป้าหมาย
บัณฑิตที่สำเร็จการศึกษาสามารถนำความรู้ไปใช้ในการพัฒนาประเทศได้มากกว่าร้อยละ
50 และโครงการฯ มีส่วนร่วมในการแก้ปัญหาการขาดแคลนอัตรากำลังคนด้าน
วิศวกรรมศาสตร์และเทคโนโลยี

¹ คณะศึกษาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

² สำนักงานคณะกรรมการการศึกษาแห่งชาติ

คำนำ

สืบเนื่องจากประเทศไทยประสบผลสำเร็จในการพัฒนาเศรษฐกิจของประเทศเป็นอย่างมาก และส่งผลให้มีการขยายตัวทางโครงสร้างพื้นฐาน เพื่อรองรับความเจริญเติบโต ในภาคอุตสาหกรรม การเกษตร การขนส่งคมนาคม ตลอดจนการจัดการทรัพยากรและสิ่งแวดล้อมให้เหมาะสม ทำให้เกิดภาวะการขาดแคลนกำลังคนทางวิศวกรรม และเทคโนโลยีอย่างมากมายังได้กล่าวแล้ว แม้รัฐบาลไทยโดยหน่วยงานหลักที่รับผิดชอบในการพัฒนาบุคลากร จะได้จัดทำแผนเร่งรัดและเพิ่มการผลิตบุคลากร โดยการขยายงานของสถาบันการศึกษาของรัฐบาล และส่งเสริมให้สถาบันอุดมศึกษา เอกชน เข้ามามีส่วนร่วมในการผลิตบุคลากรในสาขาที่ขาดแคลนอย่างต่อเนื่องก็ตาม แต่การผลิตวิศวกรและนักเทคโนโลยีเท่าที่มหาวิทยาลัยต่าง ๆ ได้เร่งรัดดำเนินการเท่าที่กล่าวมานั้น ส่วนใหญ่จะเป็นการให้การศึกษาระดับปริญญาตรีให้มีจำนวนมากที่สุด เพื่อมีบุคลากรที่สามารถปฏิบัติงานได้ในระดับพื้นฐานไปพลางก่อน และจากกรณีวิเคราะห์ระบบการศึกษาระดับมหาวิทยาลัย หรือสถาบันอุดมศึกษาในประเทศไทย รวมทั้งการศึกษาจากการประกอบการทางวิศวกรรมและเทคโนโลยีพบว่า

1. การดำเนินการสอนนักศึกษาในระดับปริญญาตรี จำเป็นต้องใช้คณาจารย์ที่มีคุณวุฒิสูงกว่าระดับปริญญาตรี เพื่อให้สามารถถ่ายทอดความรู้ที่ทันสมัย และทันต่อความก้าวหน้าทางเทคโนโลยี ส่วนใหญ่จะเป็นผู้ที่จบการศึกษาในระดับปริญญาโท และถ้าจะให้ดียิ่งขึ้นควรจบการศึกษาในระดับปริญญาเอก

2. ในการประกอบวิชาชีพทางวิศวกรรมและเทคโนโลยีนั้น ผู้ที่จบการศึกษาในระดับปริญญาตรี มีขีดความสามารถดำเนินการได้ในระดับจำกัด การปฏิบัติงานจะทำได้เฉพาะงานที่อยู่ในวงแคบ เฉพาะเรื่องต้องใช้เวลาหาความชำนาญหลายปี (ไม่ต่ำกว่า 10 ปี) จึงจะสามารถรับผิดชอบที่ต้องใช้เทคโนโลยีระดับสูง และการจัดการที่มีประสิทธิภาพได้ ฉะนั้นในการดำเนินงานทางวิศวกรรมและเทคโนโลยีในปัจจุบันและอนาคตอันใกล้นี้ จำเป็นต้องอาศัยวิศวกรและนักเทคโนโลยีที่มีการศึกษาระดับปริญญาโทและปริญญาเอกเพิ่มมากขึ้น

ด้วยเหตุนี้ โครงการพัฒนาบุคลากรทางวิศวกรรมศาสตร์ และเทคโนโลยีของสถาบันเทคโนโลยีแห่งเอเชีย จึงเกิดขึ้นเพื่อให้ความร่วมมือแก่รัฐบาลไทย ในการช่วยผลิตและสนองความต้องการบุคลากรในระดับสูงในการพัฒนาประเทศดังกล่าว ซึ่งโครงการพัฒนาบุคลากรฯ ก็ได้เริ่มดำเนินการรับนักศึกษามาตั้งแต่เดือนพฤษภาคม 2535 เป็นต้นมาจนถึงปัจจุบัน และจะสิ้นสุดโครงการฯ ในปีงบประมาณ 2539 สำนักงานคณะกรรมการการศึกษาแห่งชาติ ได้พิจารณาเห็นว่าโครงการฯนี้มีประโยชน์ต่อการผลิตและพัฒนาบุคลากรที่ขาดแคลนในสาขาวิศวกรรมศาสตร์

วิทยาศาสตร์ และเทคโนโลยีให้แก่ประเทศไทยโดยตรง จึงจำเป็นจะต้องมีการติดตามและประเมินโครงการฯ นี้ เพื่อที่จะนำผลการติดตามและประเมินจากการดำเนินโครงการฯ มาได้ระยะหนึ่ง สำหรับการเตรียมการเสนอต่อคณะกรรมการการศึกษาแห่งชาติ และคณะรัฐมนตรี เพื่อพิจารณาอนุมัติให้มีการดำเนินการโครงการพัฒนาบุคลากรฯ นี้ต่อในระยะที่ 2 (พ.ศ. 2540 - 2544) ต่อไป หรือจะยุติโครงการฯ แล้วแต่กรณี

วัตถุประสงค์ของการประเมินโครงการฯ

เพื่อให้ทราบผลการดำเนินโครงการพัฒนาบุคลากรฯ ของสถาบันเทคโนโลยีแห่งเอเชีย ที่ได้ดำเนินไปในช่วงระยะเวลาที่ผ่านมา ในการติดตามและประเมินผลการดำเนินโครงการฯ ดังกล่าว จึงกำหนดวัตถุประสงค์ของการประเมินไว้ดังนี้

1. เพื่อติดตาม และประเมินการดำเนินงาน ของโครงการพัฒนาบุคลากรทางวิศวกรรมศาสตร์และเทคโนโลยี ตามเป้าหมายในการผลิตบุคลากรในสาขาดังกล่าว คือ

ระดับปริญญาโท 45 คน/ปี

ระดับปริญญาเอก 5 คน/ปี

2. เพื่อติดตามบัณฑิตที่สำเร็จการศึกษาไปแล้ว ว่าได้นำความรู้ความสามารถไปใช้ให้เกิดประโยชน์ในการพัฒนาประเทศเพียงไร

3. เพื่อประเมินผลโครงการฯ ว่ามีส่วนร่วมในการแก้ไขปัญหาในการขาดแคลนอัตรากำลังคนด้านวิศวกรรมศาสตร์และเทคโนโลยี ทั้งในส่วนราชการ สถาบันการศึกษา และธุรกิจภาคเอกชน ได้มากน้อยเพียงใด

ขอบเขตของการประเมินโครงการฯ

ในการติดตามและประเมินผลการดำเนินงานในโครงการพัฒนาบุคลากรฯ ของสถาบันเทคโนโลยีแห่งเอเชียนี้ได้กำหนดขอบเขตของการติดตามและประเมินฯ ไว้ดังนี้

1. การผลิตบุคลากรในระดับปริญญาโทและปริญญาเอกของสถาบันเทคโนโลยีแห่งเอเชียนี้จะติดตามผลการผลิตบุคลากรเฉพาะในโครงการพัฒนาบุคลากรฯ ตามที่ได้ตกลงในบันทึกความเข้าใจ (MOU) ระหว่างรัฐบาลไทยและสถาบันเทคโนโลยีในช่วงพฤษภาคม 2535 - พฤษภาคม 2539

2. ในการติดตามผลการนำความรู้ความสามารถ ไปใช้ให้เกิดประโยชน์ในการพัฒนาประเทศนั้น จะติดตามสัมภาษณ์เฉพาะบัณฑิตที่สำเร็จการศึกษาไปแล้ว ในช่วงเมษายน 2536 ถึง

ธันวาคม 2538 และติดต่อสัมภาษณ์หัวหน้างาน หรือผู้บังคับบัญชาของบัณฑิตผู้สำเร็จการศึกษานั้นทำงานอยู่ทั้งในกรุงเทพฯ และจังหวัดต่าง ๆ ยกเว้นผู้ที่ไปทำงานในต่างประเทศ

3. ในการประเมินผลของโครงการฯ ที่เกี่ยวกับการมีส่วนร่วมช่วยแก้ปัญหาความขาดแคลนอัตรากำลังคนฯ นี้ จะเปรียบเทียบเฉพาะในภาพรวมโดยเปรียบเทียบกับเป้าหมายการประมาณการทั้งในกระบวนการผลิตกำลังคนระดับสูงกว่าปริญญาตรีและความต้องการกำลังคนระดับสูงกว่าปริญญาตรีโดยเฉลี่ยต่อปี ในสาขาวิศวกรรมศาสตร์ วิทยาศาสตร์ และเทคโนโลยี จากรายงานฉบับสมบูรณ์ของโครงการ “ยุทธศาสตร์การพัฒนากำลังคนทางเทคโนโลยีเพื่อการพัฒนาเศรษฐกิจและความสามารถในการแข่งขัน” ในฝ่ายวิจัยการพัฒนา วิทยาศาสตร์และเทคโนโลยี สถาบันวิจัยเพื่อการพัฒนาประเทศไทย ที่เสนอกระทรวงวิทยาศาสตร์และเทคโนโลยีและสิ่งแวดล้อม (กุมภาพันธ์ 2538) และจากรายงานการวิจัยของสถาบันวิจัยเพื่อการพัฒนาเรื่อง Projects of Manpower Usage in Industry and Service (TDRI), 1992

วิธีการประเมิน

แหล่งข้อมูล

ข้อมูลที่ใช้ในการประเมินผล มี 4 ส่วน คือ

1. ข้อมูลจากเอกสารงานทะเบียนเกี่ยวข้องกับบัณฑิตของโครงการฯ ซึ่งเก็บบันทึกไว้ที่งานทะเบียนของสถาบันเทคโนโลยีแห่งเอเชีย
2. ข้อมูลจากการสัมภาษณ์และตอบแบบสอบถามของบัณฑิตของโครงการที่จบการศึกษาในช่วงปีงบประมาณ 2535 - 2539
3. ข้อมูลจากการสัมภาษณ์ผู้บังคับบัญชาของบัณฑิตของโครงการ
4. ข้อมูลจากการสัมภาษณ์ผู้ที่เกี่ยวข้องในกระบวนการรับนักศึกษา ตามโครงการพัฒนาบุคลากรทางวิศวกรรมศาสตร์ และเทคโนโลยี

เครื่องมือที่ใช้ในการรวบรวมข้อมูล

เครื่องมือที่ใช้ในการรวบรวมข้อมูลประกอบด้วย

1. แบบวิเคราะห์ข้อมูลจากเอกสารและสิ่งตีพิมพ์เกี่ยวกับจำนวนนักศึกษาที่เข้าเรียนในโครงการในแต่ละปีงบประมาณ และจำนวนบัณฑิตที่จบการศึกษา แยกตามสาขาวิชาต่าง ๆ ในแต่ละระดับปริญญาโทหรือเอก
2. แบบสัมภาษณ์และแบบสอบถามบัณฑิตของโครงการฯ
3. แบบสัมภาษณ์ผู้บังคับบัญชาของบัณฑิต

4. แบบสัมภาษณ์ผู้ที่เกี่ยวข้องในกระบวนการรับนักศึกษา ตามโครงการพัฒนาบุคลากรทางวิศวกรรมศาสตร์ และเทคโนโลยี

เกณฑ์ในการประเมิน

1. นโยบายและเป้าหมายในการรับและการผลิตบัณฑิต
2. ระยะเวลาในการศึกษาของบัณฑิตตามโครงการ
3. จำนวนและร้อยละของบัณฑิตที่สามารถนำความรู้ไปใช้ประโยชน์ในด้านวิศวกรรมศาสตร์และเทคโนโลยี เพื่อการพัฒนาประเทศ

รูปแบบของการประเมิน

การประเมินผลความร่วมมือระหว่างรัฐบาลไทยและสถาบันเทคโนโลยีแห่งเอเชีย ในโครงการพัฒนาบุคลากรทางวิศวกรรมศาสตร์และเทคโนโลยี เป็นการประเมินภายหลังจากโครงการได้ดำเนินงานมาครบ 4 ปี (13 ภาคการศึกษา) เพื่อนำผลการประเมินมาใช้ในการตัดสินใจเกี่ยวกับการปรับปรุงการดำเนินโครงการ หรือการพิจารณาดำเนินงานโครงการต่อไปหรือไม่ มีรูปแบบการประเมินโดยการนำผลการดำเนินงานของโครงการ ในช่วงปีพุทธศักราช 2535 - 2539 เทียบกับเกณฑ์ คือ นโยบายและเป้าหมายในการรับและการผลิตบัณฑิตในแต่ละระดับในแต่ละปี รวมทั้งเกณฑ์ระยะเวลาในการศึกษาของบัณฑิตตามโครงการและจำนวนร้อยละของบัณฑิตที่สามารถนำความรู้ไปใช้ประโยชน์ในด้านวิศวกรรมศาสตร์และเทคโนโลยีเพื่อการพัฒนาประเทศ โดยมีขอบเขตของการประเมินเพื่อติดตามผลบัณฑิตของโครงการที่สำเร็จการศึกษาในช่วงเมษายน 2536 - ธันวาคม 2538

การรวบรวมข้อมูล

1. ผู้ประเมินติดตามรายชื่อบัณฑิตของโครงการในช่วง พฤษภาคม 2535 - พฤษภาคม 2539 โดยส่งแบบสอบถามทางไปรษณีย์ถึงบัณฑิตทุกคนที่สำเร็จการศึกษาในช่วงเมษายน 2536 - ธันวาคม 2538 พร้อมทั้งให้บัณฑิตระบุรายชื่อผู้บังคับบัญชาของบัณฑิต และที่อยู่ของผู้บังคับบัญชา เพื่อผู้ประเมินจะได้ติดตามเพื่อดำเนินการสัมภาษณ์ต่อไป
2. ผู้ประเมินติดต่อผู้บังคับบัญชาของบัณฑิต ตามที่บัณฑิตระบุโทรศัพท์เพื่อนัดหมายเวลาและสถานที่ในการสัมภาษณ์ และดำเนินการสัมภาษณ์
3. ผู้ประเมินรวบรวมข้อมูลที่เกี่ยวข้องกับบัณฑิตของโครงการจากฝ่ายทะเบียนของสถาบันเทคโนโลยีแห่งเอเชีย

4. ผู้ประเมินรวบรวมข้อมูลเกี่ยวกับกระบวนการรับนักศึกษา โดยการสัมภาษณ์ผู้ที่เกี่ยวข้อง ในกระบวนการรับนักศึกษา จำนวน 6 คน

การวิเคราะห์ข้อมูล

1. คำนวณค่าความถี่และร้อยละของบัณฑิตในแต่ละปี แยกตามระดับและสาขาวิชา เปรียบเทียบกับจำนวนเป้าหมายตามแผนที่กำหนดของโครงการฯ
2. วิเคราะห์เนื้อหาคำตอบของบัณฑิตเกี่ยวกับจุดเด่น จุดด้อย ของการดำเนินงานของโครงการฯ
3. วิเคราะห์เนื้อหาคำตอบของผู้บังคับบัญชาเกี่ยวกับความรู้ความสามารถ และคุณลักษณะของบัณฑิต
4. วิเคราะห์เนื้อหาคำตอบของผู้ที่เกี่ยวข้องกับกระบวนการรับนักศึกษา

ผลการประเมินโครงการ

1. การผลิตบุคลากรตามเป้าหมาย

1.1 จำนวนนักศึกษาที่รับในระดับปริญญาโท โดยเฉลี่ยเท่ากับเป้าหมาย แต่จำนวนนักศึกษาที่รับในระดับปริญญาเอกสูงกว่าเป้าหมายเท่าตัว เมื่อพิจารณาถึงความสอดคล้องของจำนวนรับนักศึกษา จำแนกตามหน่วยงานของนักศึกษาพบว่าในส่วนของทุนจากรัฐบาลไทย การรับนักศึกษามีความสอดคล้องกับลำดับความสำคัญของเกณฑ์การพิจารณาการคัดเลือกที่กำหนดไว้เพียงบางส่วน คือมีนักศึกษาที่มาจากหน่วยราชการมีจำนวนมากที่สุด ซึ่งตามเกณฑ์การพิจารณาการคัดเลือกได้กำหนดไว้ว่า ผู้สมัครที่มาจากหน่วยราชการจะได้รับการพิจารณาเป็นอันดับแรก แต่ปรากฏว่าจำนวนนักศึกษาที่มาจากมหาวิทยาลัยของรัฐในส่วนภูมิภาคจัดอยู่ในอันดับ 4 ซึ่งผู้สมัครกลุ่มนี้ตามเกณฑ์การพิจารณามีความสำคัญเป็นอันดับสอง

ในส่วนของทุนจากสถาบันฯ ปรากฏว่านักศึกษาส่วนใหญ่มาจากองค์กรเอกชน ซึ่งไม่สอดคล้องกับลำดับความสำคัญของเกณฑ์การพิจารณาการคัดเลือกที่กำหนดให้ผู้สมัครที่มาจากหน่วยราชการจะได้รับการพิจารณาเป็นอันดับแรก

เมื่อพิจารณาจำนวนนักศึกษาที่รับเข้าศึกษาที่สถาบันฯ จำแนกตามสาขาวิชาปรากฏว่า มีนักศึกษาเข้าศึกษาในทุกสาขาวิชา สาขาที่มีนักศึกษาของโครงการฯ เข้าศึกษามากที่สุดคือ สาขาวิศวกรรมอุตสาหการ และวิศวกรรมการเกษตรและอาหาร รองลงมาคือสาขาวิทยาการคอมพิวเตอร์ การวางแผนพัฒนาชุมชนและท้องถิ่น วิศวกรรมสิ่งแวดล้อม วิศวกรรมการผลิต วิศวกรรมโครงสร้างและการก่อสร้าง นอกนั้นมีจำนวนไม่เกินสาขาละ 10 คน ดังแสดงในตารางที่ 1

ตารางที่ 1 จำนวนนักศึกษาที่รับและจบการศึกษาในช่วง พ.ศ. 2535-พ.ศ. 2539

สาขาวิชา	สถานภาพ	ปี พ.ศ.															รวม		
		2535			2536			2537			2538			2539					
		RTG	HRD	รวม	RTG	HRD	รวม	RTG	HRD	รวม	RTG	HRD	รวม	RTG	HRD	รวม	RTG	HRD	รวม
1. วิทยาการคอมพิวเตอร์	รับ	4	3	7	5	3	11	2	-	2	1	1	2	3	1	4	15	11	26
	จบ	-	-	-	1	3	4	3	4	7	1	-	1	NA	NA	NA	5	7	12
2. การวางแผนพัฒนาชนบทและภูมิภาค	รับ	5	1	6	4	3	7	2	-	2	1	1	2	1	-	1	13	5	18
	จบ	-	-	-	2	1	3	2	3	5	-	-	-	NA	NA	NA	4	4	8
3. วิศวกรรมอุตสาหกรรม	รับ	7	4	11	5	5	10	2	2	4	3	5	8	1	1	2	18	17	35
	จบ	-	-	-	5	4	9	5	2	7	1	5	6	NA	NA	NA	11	11	22
4. การจัดการและการวางแผนทรัพยากรธรรมชาติ	รับ	1	1	2	1	-	1	3	1	4	1	-	1	-	1	1	6	3	9
	จบ	-	-	-	-	1	1	1	-	1	1	1	2	NA	NA	NA	2	2	4
5. วิศวกรรมสิ่งแวดล้อม	รับ	4	1	5	-	-	-	2	4	6	3	3	6	-	-	-	9	8	17
	จบ	-	-	-	-	-	-	3	1	4	-	4	4	NA	NA	NA	3	5	8
6. วิศวกรรมการขนส่ง	รับ	4	3	7	3	-	3	2	2	4	1	-	1	-	-	-	10	5	15
	จบ	-	-	-	1	-	1	2	2	4	3	-	3	NA	NA	NA	6	2	8
7. วิศวกรรมโครงสร้างและการก่อสร้าง	รับ	3	4	7	1	-	1	1	2	3	-	-	-	-	-	-	5	6	11
	จบ	-	-	-	-	-	-	3	1	7	1	-	1	NA	NA	NA	4	4	3
8. การจัดการเทคโนโลยี	รับ	2	2	4	3	-	3	2	2	3	2	1	3	-	-	-	9	4	13
	จบ	-	-	-	-	-	-	2	6	4	3	-	3	NA	NA	NA	5	2	7
9. วิศวกรรมการเกษตรและอาหาร	รับ	4	-	4	5	3	8	7	3	13	5	5	10	-	-	-	21	14	35
	จบ	-	-	-	-	-	-	2	-	5	6	5	-	-	-	-	-	6	-
10. เทคโนโลยีพลังงาน	รับ	-	-	-	1	3	4	2	-	2	-	1	1	-	2	2	3	6	9
	จบ	-	-	-	-	-	-	-	3	3	2	-	2	NA	NA	NA	2	3	5
11. วิศวกรรมคอมพิวเตอร์	รับ	-	-	-	1	-	1	-	1	1	-	2	2	2	-	2	3	3	6
	จบ	-	-	-	-	-	-	1	-	1	-	1	1	NA	NA	NA	1	1	2

ตารางที่ 1 (ต่อ)

สาขาวิชา	สถานภาพ	ปี พ.ศ.															รวม		
		2535			2536			2537			2538			2539					
		RTG	HRD	รวม	RTG	HRD	รวม	RTG	HRD	รวม	RTG	HRD	รวม	RTG	HRD	รวม	RTG	HRD	รวม
12. วิศวกรรมแหล่งน้ำ	รับ	-	-	-	1	-	1	2	-	2	-	-	-	-	-	-	3	-	3
	จบ	-	-	-	-	-	-	-	-	-	-	-	-	NA	NA	NA	-	-	-
13. Spring Program	รับ	-	-	-	1	-	1	-	-	-	-	-	-	-	-	-	1	-	1
	จบ	-	-	-	-	-	-	-	-	-	1	-	1	NA	NA	NA	1	-	1
14. วิศวกรรมชลประทานและจัดการ	รับ	-	-	-	-	-	-	2	-	2	-	-	-	-	-	-	2	-	2
	จบ	-	-	-	-	-	-	-	-	-	-	-	-	NA	NA	NA	-	-	-
15. การจัดการสารสนเทศ	รับ	-	-	-	-	-	-	-	-	-	2	-	2	1	-	1	3	-	3
	จบ	-	-	-	-	-	-	-	-	-	-	-	-	NA	NA	NA	-	-	-
16. วิศวกรรมระบบการผลิต	รับ	-	-	-	-	-	-	-	-	-	-	-2	2	-	-	-	-	2	2
	จบ	-	-	-	-	-	-	-	-	-	-	-	-	NA	NA	NA	-	-	-
17. ธรณีวิศวกรรมและธรณีวิทยาประยุกต์	รับ	-	-	-	-	-	-	-	-	-	2	1	3	-	-	-	2	1	3
	จบ	-	-	-	-	-	-	-	-	-	-	-	-	NA	NA	NA	-	-	-
18. วิศวกรรมปฐพี	รับ	-	-	-	-	-	-	-	-	-	1	1	2	-	-	-	1	1	2
	จบ	-	-	-	-	-	-	-	-	-	-	-	-	NA	NA	NA	-	-	-
19. การวางแผนและจัดการโครงสร้างพื้นฐาน	รับ	-	-	-	-	-	-	-	-	-	1	1	2	-	-	-	1	1	2
	จบ	-	-	-	-	-	-	-	-	-	-	-	-	NA	NA	NA	-	-	-
20. การพัฒนาแหล่งน้ำ	รับ	-	-	-	-	-	-	-	-	-	1	-	1	-	-	-	1	-	1
	จบ	-	-	-	-	-	-	-	-	-	-	-	-	2	NA	NA	-	-	-
21. เทคโนโลยีหลังการเก็บเกี่ยวและวิศวกรรมอาหาร	รับ	-	-	-	-	-	-	-	-	-	-	-	-	2	-	2	2	-	2
	จบ	-	-	-	-	-	-	-	-	-	-	-	-	NA	NA	NA	-	-	-
22. เทคโนโลยีการเพาะเลี้ยงสัตว์น้ำ	รับ	-	-	-	-	-	-	-	-	-	-	-	-	2	4	6	2	4	6
	จบ	-	-	-	-	-	-	-	-	-	-	-	-	NA	NA	NA	-	-	-

ตารางที่ 1 (ต่อ)

สาขาวิชา	สถานภาพ	ปี พ.ศ.															รวม		
		2535			2536			2537			2538			2539					
		RTG	HRD	รวม	RTG	HRD	รวม	RTG	HRD	รวม	RTG	HRD	รวม	RTG	HRD	รวม	RTG	HRD	รวม
23. วิศวกรรมการประปาและน้ำเสีย	รับ	-	-	-	-	-	-	-	-	-	-	-	-	1	-	1	1	-	1
	จบ	-	-	-	-	-	-	-	-	-	-	-	-	NA	NA	NA	-	-	-
24. เทคโนโลยีขบวนการชีวภาพ	รับ	-	-	-	-	-	-	-	-	-	-	-	-	1	2	3	1	2	3
	จบ	-	-	-	-	-	-	-	-	-	-	-	-	NA	NA	NA	-	-	-
25. เทคโนโลยีและการจัดการสิ่งแวดล้อม	รับ	-	-	-	-	-	-	-	-	-	-	-	-	1	4	5	1	4	5
	จบ	-	-	-	-	-	-	-	-	-	-	-	-	NA	NA	NA	-	-	-
26. การสำรวจระยะไกลและสารสนเทศภูมิศาสตร์	รับ	-	-	-	-	-	-	-	-	-	-	-	-	1	1	2	1	1	2
	จบ	-	-	-	-	-	-	-	-	-	-	-	-	NA	NA	NA	-	-	-
27. วิศวกรรมการก่อสร้างและการจัดการ	รับ	-	-	-	-	-	-	-	-	-	-	2	2	-	-	-	-	2	2
	จบ	-	-	-	-	-	-	-	-	-	-	-	-	NA	NA	NA	-	-	-
28. วิศวกรรมโครงสร้าง	รับ	-	-	-	-	-	-	-	-	-	2	3	5	-	-	-	2	3	5
	จบ	-	-	-	-	-	-	-	-	-	-	-	-	NA	NA	NA	-	-	-
รวม	รับ	34	19	53	31	20	51	29	19	48	26	29	55	16	16	32	136	103	239
	จบ	-	-	-	9	9	18	24	24	48	19	16	35	NA	NA	NA	52	49	101

1.2 ในช่วงพฤษภาคม 2535 - ธันวาคม 2538 มีบัณฑิตที่สำเร็จจากโครงการไปแล้วจำนวนทั้งสิ้น 101 คน เป็นนักศึกษาในระดับปริญญาโท 99 คน ปริญญาเอก 2 คน นักศึกษาปริญญาโทส่วนใหญ่ใช้เวลาในการศึกษา 5 ภาคการศึกษา สำหรับนักศึกษาระดับปริญญาเอกใช้เวลาศึกษา 9 ภาคการศึกษา 1 คน และ 11 ภาคการศึกษา 1 คน นอกนั้นยังไม่สำเร็จการศึกษา เป็นที่น่าสังเกตว่านักศึกษาระดับปริญญาเอกส่วนใหญ่ใช้เวลามากกว่า 9 ภาคการศึกษา

1.3 จากจำนวนนักศึกษาที่รับเข้ามาในช่วง พฤษภาคม 2535 - พฤษภาคม 2539 จำนวนรวมทั้งสิ้น 239 คน มีผู้ลาออกกลางคันเพียง 7 คน เท่านั้น (ร้อยละ 2.93) ซึ่งสาเหตุของการลาออกกลางคันนั้น เนื่องมาจากปัญหาเกี่ยวกับงานประจำที่นักศึกษารับผิดชอบ ปัญหาเกี่ยวกับครอบครัว ปัญหาเกี่ยวกับสาขาวิชาที่เรียน ปัญหาเกี่ยวกับการปรับตัวของนักศึกษาและปัญหาสืบเนื่องมาจากการได้รับแจ้งผลทุนการศึกษาที่ค่อนข้างล่าช้า

1.4 กระบวนการรับนักศึกษาของโครงการฯ นั้น ในส่วนของทุนจากรัฐบาลไทยผู้มีส่วนร่วมในกระบวนการรับนักศึกษาคือ สำนักงานคณะกรรมการข้าราชการพลเรือน สำนักงานคณะกรรมการการศึกษาแห่งชาติ และสถาบันฯ แต่ในส่วนของสถาบันฯ ในวงเงิน 10 ล้านบาท สถาบันฯ จะเป็นผู้พิจารณาเองทั้งหมด ซึ่งสถาบันฯ จะจัดสรรในหลายลักษณะดังนี้ ให้ทุนเต็มเวลาให้ทุนเป็นบางส่วน ให้ในลักษณะลดหย่อนค่าเล่าเรียน ให้ในลักษณะชดเชยเพิ่มเติมให้กับวงเงินในส่วนที่รัฐบาลไทยรับผิดชอบ เช่น ในช่วงที่มีการขึ้นค่าเล่าเรียน ปัญหาที่เกิดขึ้นในกระบวนการรับนักศึกษาก็คือรัฐบาลไทยไม่ได้ระบุให้ชัดเจนว่าต้องการบุคลากรในสาขาใดมากกว่ากัน ดังนั้นสถาบันฯ จะพิจารณาบนพื้นฐานของอัตรากำลังอาจารย์ที่มีอยู่ ซึ่งอาจจะไม่ตรงกับความต้องการของรัฐบาลไทย และจากการสอบถามความคิดเห็นของนักศึกษา พบว่าปัญหาที่เกิดขึ้นในกระบวนการรับนักศึกษา คือ การแจ้งผลให้ผู้สมัครทราบค่อนข้างล่าช้ามาก ซึ่งมีผลกระทบต่อนักศึกษาบางคนที่ไม่สามารถเรียนได้อย่างเต็มที่เนื่องจากยังต้องกลับไปทำงาน เพราะหน่วยงานไม่สามารถหาคนมาทดแทนได้ทัน

2. การนำความรู้ความสามารถไปใช้ประโยชน์ ในการพัฒนาประเทศ ผู้ประเมินได้สอบถามความคิดเห็นของบัณฑิตที่สำเร็จจากโครงการ และผู้บังคับบัญชาของบัณฑิต ปรากฏผลดังนี้

2.1 ความคิดเห็นของบัณฑิต จากการสอบถามความคิดเห็นของบัณฑิตที่สำเร็จจากโครงการ เกี่ยวกับปัญหาที่เกิดขึ้นในระหว่างศึกษาอยู่ กระบวนการจัดการเรียนการสอนของสถาบันฯ การนำความรู้ความสามารถไปใช้ประโยชน์และความคิดเห็นเกี่ยวกับโครงการพบว่า

2.1.1 ผู้ที่ได้ทุนจากรัฐบาลไทยและจากสถาบันฯ ระบุว่าไม่มีปัญหาเกี่ยวกับเงินทุนไม่มีปัญหาเกี่ยวกับการเรียน และไม่มีปัญหาเกี่ยวกับการทำวิทยานิพนธ์ มีจำนวนมากที่สุด แต่

เป็นที่น่าสังเกตว่ามีบัณฑิตจำนวนหนึ่ง (ร้อยละ 19 ถึงร้อยละ 44) ที่ระบุว่ามีปัญหาทุนไม่เพียงพอ มีปัญหาเรียนไม่ทัน มีปัญหาด้านภาษา อุปกรณ์ที่ใช้ในการเรียนการสอน เช่น คอมพิวเตอร์ มีจำนวนไม่พอกับจำนวนนักศึกษา และค่าใช้จ่ายในการทำวิทยานิพนธ์ไม่พอ

2.1.2 บัณฑิตส่วนใหญ่ (ร้อยละ 77.77) ในส่วนของทุนจากรัฐบาลไทย และ (ร้อยละ 63.63) ในส่วนของทุนของสถาบันฯ ระบุว่าได้นำความรู้และประสบการณ์ที่ได้รับจากสถาบันฯ ไปใช้ในการปฏิบัติงานมากกว่าร้อยละ 50 ของงานที่ทำ แต่ต้องไปฝึกปฏิบัติเพิ่มเติมเพื่อฝึกทักษะและความชำนาญให้มากขึ้น นอกจากนั้นต้องปรับตัวเองให้ทันกับความเปลี่ยนแปลงของโลกและเทคโนโลยี สำหรับประสิทธิภาพในการทำงานหลังจากจบการศึกษา ส่วนใหญ่ของบัณฑิตระบุว่ามีประสิทธิภาพดีขึ้น เนื่องจากมีความมั่นใจมากขึ้น กล้าตัดสินใจมากขึ้น สามารถทำงานได้อย่างเป็นระบบมากขึ้น มีความขยันมากขึ้น หาแนวทางแก้ปัญหาได้ดีกว่าเดิม มีความเข้าใจรายละเอียดด้านวิชาการมากขึ้น มีความรู้เกี่ยวกับคอมพิวเตอร์มากขึ้น มีความกระตือรือร้นในการค้นคว้าวารสารและเอกสารจากต่างประเทศเพิ่มขึ้น ซึ่งเป็นประโยชน์ในการเพิ่มพูนความรู้ คิดได้กว้างและลึกมากขึ้น มีความสามารถในการวิจัยดีขึ้น มีกระบวนการคิดที่เป็นระบบมากขึ้น ซึ่งเป็นประโยชน์ต่อการทำงานเป็นอย่างมาก ดังแสดงในตารางที่ 2

2.1.3 บัณฑิตทุกคนจากทุนรัฐบาลไทยและทุนของสถาบันฯ มีความเห็นว่าโครงการนี้ควรจะดำเนินต่อไป ซึ่งจะเป็นการช่วยเพิ่มบุคลากรทางด้านวิทยาศาสตร์และเทคโนโลยีช่วยให้ผู้ที่มีทุนการศึกษาน้อยได้มีโอกาสศึกษาต่อ เป็นการเปิดโอกาสให้คนไทยได้เข้าศึกษาในสถาบันนานาชาติ โดยเฉพาะผู้ที่ไม่สามารถไปศึกษาต่อต่างประเทศ ด้วยเหตุผลในด้านของการขาดแคลนทุนทรัพย์ หรือภาระครอบครัว และการเรียนการสอนที่สถาบันฯ เป็นที่ยอมรับในระดับนานาชาติ ซึ่งผู้ที่เข้าศึกษาสามารถนำมาใช้เป็นแบบอย่างที่ดีต่อไปได้

2.1.4 บัณฑิตส่วนใหญ่มีความเห็นว่าโครงการควรปรับปรุงกระบวนการรับนักศึกษาในประเด็นดังต่อไปนี้คือ ควรเพิ่มจำนวนทุนให้มากขึ้น ควรคัดเลือกผู้ที่มีความสามารถพอสมควรทั้งด้านวิชาการและภาษาอังกฤษ ควรปรับปรุงการประชาสัมพันธ์ ควรมีหลักเกณฑ์ที่แน่นอนในการคัดเลือก ควรมีการสอบคัดเลือก และจำนวนนักศึกษาที่รับควรมีสัดส่วนที่พอเหมาะกับจำนวนอาจารย์และเครื่องมืออุปกรณ์ต่าง ๆ ในประเด็นของเงินเดือนและทุนที่ให้นั้น บัณฑิตมีความเห็นว่าควรเพิ่มเงินทุนให้เพียงพอกับค่าใช้จ่ายในการทำวิทยานิพนธ์และค่าครองชีพ เงินไขการให้ทุนควรจะรัดกุมกว่านี้ ควรให้ทุนตลอดการศึกษา และควรให้ใช้ทุนโดยรับราชการ สำหรับประเด็นของหลักสูตร บัณฑิตที่สำเร็จจากโครงการทั้งสองกลุ่มมีความเห็นว่าเหมาะสมดีแล้ว มีจำนวนมากที่สุด ประเด็นที่ควรปรับปรุง คือควรปรับหลักสูตรให้ทันสมัยตามความก้าวหน้าของเทคโนโลยีอย่าง

สม่ำเสมอ ควรให้นักศึกษาไปสัมผัสกับหน่วยงานของรัฐและเอกชนต่าง ๆ เพื่อให้เกิดความคิดที่จะแก้ปัญหาและเพิ่มประสบการณ์

ตารางที่ 2 การนำความรู้และประสบการณ์ที่ได้รับจากสถาบันเทคโนโลยีแห่งเอเชียมาใช้ในการปฏิบัติงาน

ปัญหา	RTG (n = 36)		HRD (n = 33)		รวม (n = 69)	
	จำนวน	ร้อยละ	จำนวน	ร้อยละ	จำนวน	ร้อยละ
ร้อยละของงานที่ใช้ความรู้และประสบการณ์ที่ได้รับจากสถาบันฯ						
ไม่สามารถนำมาใช้ได้เลย	1	2.78	–	–	1	1.45
ต่ำกว่า 50	6	16.67	9	27.27	15	21.74
50-70	16	44.44	13	39.39	29	42.03
71-90	4	11.11	8	24.24	12	17.39
91-100	8	22.22	–	–	8	11.59
ไม่ตอบ	1	2.78	3	9.09	4	5.80
การฝึกปฏิบัติเพิ่มเติมหลังจากการศึกษาจากสถาบันฯ						
ไม่ต้องฝึกปฏิบัติเพิ่มเติม	7	19.44	5	15.15	12	17.39
ต้องฝึกปฏิบัติเพิ่มเติม	24	66.67	23	69.70	47	68.12
ไม่ตอบ	5	13.89	5	15.15	10	14.49
ประสิทธิภาพในการทำงานหลังจากจบการศึกษาจากสถาบันฯ						
ดีขึ้นมากและสามารถทำงาน- อย่างเป็นระบบ	2	5.55	6	18.18	8	11.59
ดีขึ้น	33	91.67	25	75.76	58	84.06
ไม่แตกต่าง	–	–	1	3.03	1	1.45
ไม่ตอบ	1	2.78	1	3.03	2	2.90

2.2 ความคิดเห็นของผู้บังคับบัญชา

2.2.1 บัณฑิตที่กลับมาทำงานในหน่วยงานเดิม ผู้บังคับบัญชามีความเห็นว่าบัณฑิตมีประสิทธิภาพในการทำงานเพิ่มขึ้น คือ มีความรู้ทางวิชาการสูงขึ้นในด้านการวางแผนการทำงาน การวิเคราะห์งานคอมพิวเตอร์และภาษาอังกฤษ มีเจตคติที่ดีต่อการทำงาน มีความกระตือรือร้นและตั้งใจทำงาน ไม่เกียจงาน มีความคิดสร้างสรรค์ มีความเป็นผู้นำ และมีความรับผิดชอบดีมาก ผู้บังคับบัญชามีความเห็นว่า การเปลี่ยนแปลงไปในทางที่ดีขึ้นนั้นส่วนหนึ่งมาจากการศึกษาต่อที่สถาบันฯ

2.2.2 บัณฑิตที่จบจากสถาบันฯและหน่วยงานรับเข้ามาใหม่นั้น ผู้บังคับบัญชามีความเห็นว่าเหตุผลของการรับบัณฑิตเข้ามาทำงานนั้นเนื่องจากสาขาและความสามารถของบัณฑิตที่จบมานั้นตรงตามความต้องการของหน่วยงาน และบุคลิกภาพที่เหมาะสม

2.2.3 เมื่อเปรียบเทียบความสามารถของบัณฑิตที่จบจากสถาบันฯ กับบัณฑิตที่จบจากสถาบันอื่นที่มาทำงานในหน่วยงานเดียวกัน ผู้บังคับบัญชามีความเห็นว่าบัณฑิตที่จบจากสถาบันฯมีประสิทธิภาพและเจตคติในการทำงานดีกว่า คือ มีความสามารถทางคอมพิวเตอร์และภาษาอังกฤษดีกว่า มีความกระตือรือร้นในการทำงานและสู้งาน

2.2.4 ผู้บังคับบัญชามีความเห็นว่าทั้งภาครัฐและภาคเอกชน ยังมีความต้องการบุคลากรที่จบทางด้านวิศวกรรมศาสตร์และเทคโนโลยีเพิ่มเติมในสาขาต่าง ๆ อีกมาก สาขาที่มีความต้องการ ได้แก่ วิศวกรรมไฟฟ้า วิศวกรรมโยธา วิศวกรรมแหล่งน้ำ วิศวกรรมชลประทาน อุตสาหกรรมเคมี วิทยาศาสตร์การอาหารและสิ่งแวดล้อม การสำรวจระยะไกล สัตว์ป่า วิศวกรรมอุตสาหการ วิศวกรรมการขนส่ง วิศวกรรมเครื่องกล วิศวกรรมคอมพิวเตอร์ และวิศวกรรมโทรคมนาคม

3. การมีส่วนร่วมในการแก้ปัญหาการขาดแคลนอัตรากำลังคน สถาบันฯมีส่วนร่วม ในการแก้ปัญหาการขาดแคลนอัตรากำลังคนทางด้านวิศวกรรมศาสตร์และเทคโนโลยี ในระดับที่สูงกว่าปริญญาตรี โดยสถาบันฯได้ผลิตบัณฑิตในระดับสูงกว่าปริญญาตรี ในสาขาวิศวกรรมศาสตร์ เฉลี่ยประมาณปีละ 21 คน ซึ่งคิดเป็นร้อยละ 3.68 ของความต้องการบุคลากรในสภาวะปกติ ร้อยละ 2.59 ของความต้องการบุคลากรในสภาวะที่มีการเปลี่ยนแปลงทางเทคโนโลยีและโครงสร้างทางเศรษฐกิจอย่างมาก และคิดเป็นร้อยละ 6.25 ของจำนวนบัณฑิตที่ผลิตได้โดยรวม สำหรับสาขาวิทยาศาสตร์และเทคโนโลยี สถาบันฯได้ผลิตบัณฑิตในระดับสูงกว่าปริญญาตรี เฉลี่ยประมาณ 13 คนต่อปี ซึ่งคิดเป็นร้อยละ 0.92 ของความต้องการบุคลากรในสภาวะปกติ ร้อยละ 0.71 ของความต้องการบุคลากรในสภาวะที่มีการเปลี่ยนแปลงทางเทคโนโลยีและโครงสร้างทางเศรษฐกิจอย่างมาก และคิดเป็นร้อยละ 3.28 ของจำนวนบัณฑิตที่ผลิตได้โดยรวม

สรุป และข้อเสนอแนะ

1. จากผลการประเมินในช่วง 13 ภาคการศึกษาแรก คณะผู้ประเมินมีความเห็นว่าโครงการพัฒนาบุคลากรทางวิศวกรรมศาสตร์และเทคโนโลยีนั้น ควรจะดำเนินการต่อไป เนื่องจาก

1.1 การผลิตบุคลากรนั้น เมื่อพิจารณาในภาพรวมสูงกว่าเป้าหมายที่กำหนดไว้ตามเป้าหมายที่กำหนดไว้ 5 ปี (15 ภาคการศึกษา) จะรับนักศึกษาระดับปริญญาโทจำนวน 225 คน ระดับปริญญาเอก 25 คน รวมทั้งสิ้น 250 คน ณ ช่วงเวลาที่ทำการประเมินสถาบันฯรับนักศึกษาไป 13 ภาคการศึกษา ซึ่งควรจะมีจำนวนนักศึกษาทั้งหมดตามเป้าหมาย 217 คน ปรากฏว่าสถาบันฯรับนักศึกษาสูงกว่าเป้าหมาย คือ รับมาทั้งสิ้น 239 คน ซึ่งจำนวนนักศึกษาระดับปริญญาโท เท่ากับเป้าหมาย (150 คน) โดยรับมา 151 คน แต่ไม่สามารถเรียนระดับปริญญาโทได้จึงให้ประกาศนียบัตรแทน 2 คน จึงมีนักศึกษาระดับปริญญาโทในช่วง 13 ภาคการศึกษา 149 คน แต่จำนวนนักศึกษาระดับปริญญาเอกสูงกว่าเป้าหมาย (17 คน) เป็นเท่าตัว โดยในช่วง 13 ภาคการศึกษา รับนักศึกษาระดับปริญญาเอก ถึง 38 คน จากจำนวนนักศึกษาทั้งหมดนั้น มีผู้ที่ลาออกกลางคันเพียง 7 คน เป็นนักศึกษาระดับปริญญาโท 3 คน และปริญญาเอก 4 คน

1.2 จากการสอบถามความคิดเห็นของบัณฑิตและผู้บังคับบัญชา ปรากฏว่า บัณฑิตสามารถนำความรู้ที่ได้รับจากสถาบันฯมากกว่าร้อยละ 50 ไปใช้ประโยชน์ในการปฏิบัติงาน และประสิทธิภาพของการปฏิบัติงานดีขึ้น มีความมั่นใจมากขึ้น กล้าตัดสินใจมากขึ้น สามารถทำงานได้อย่างเป็นระบบมากขึ้น มีความขยันมากขึ้น มีความรู้ทางด้านวิชาการมากขึ้น มีความรู้เกี่ยวกับการใช้คอมพิวเตอร์และการใช้ภาษาอังกฤษมากขึ้น มีความกระตือรือร้นในการปฏิบัติงานมากขึ้น มีความสามารถในการวิจัยดีขึ้น มีความคิดสร้างสรรค์ และมีเจตคติที่ดีต่อการทำงาน

1.3 โครงการมีส่วนร่วมในการแก้ปัญหาการขาดแคลนอัตรากำลังคนทางด้านวิศวกรรมศาสตร์และเทคโนโลยีในระดับที่สูงกว่าปริญญาตรีโดยสถาบันฯผลิตบัณฑิตในระดับที่สูงกว่าปริญญาตรีสาขาวิศวกรรมศาสตร์ เฉลี่ยประมาณปีละ 21 คน คิดเป็นร้อยละ 3.68 ของความต้องการบุคลากรในสภาวะปกติ ร้อยละ 2.59 ของความต้องการบุคลากรในสภาวะที่มีการเปลี่ยนแปลงทางเทคโนโลยีและโครงสร้างทางเศรษฐกิจอย่างมาก และคิดเป็นร้อยละ 6.25 ของจำนวนบัณฑิตที่ผลิตได้โดยรวมสำหรับสาขาวิทยาศาสตร์และเทคโนโลยี สถาบันฯได้ผลิตบัณฑิตในระดับสูงกว่าปริญญาตรี เฉลี่ยประมาณ 13 คนต่อปี ซึ่งคิดเป็นร้อยละ 0.92 ของความต้องการบุคลากรในสภาวะปกติ ร้อยละ 0.71 ของความต้องการบุคลากรในสภาวะที่มีการเปลี่ยนแปลงทางเทคโนโลยีและโครงสร้างทางเศรษฐกิจอย่างมาก และคิดเป็นร้อยละ 3.28 ของจำนวนบัณฑิตที่ผลิตได้โดยรวม

2. ข้อเสนอแนะในการปรับปรุงการดำเนินงานของโครงการ ในช่วงต่อไปนี้มีดังนี้

2.1 การประชาสัมพันธ์เกี่ยวกับโครงการ ควรจะดำเนินการให้มากกว่านี้ ซึ่งสถาบันฯ รับนักศึกษาปีละ 3 ครั้ง ควรจะมีการประชาสัมพันธ์ไปตามหน่วยงานและสื่อมวลชน ปีละ 3 ครั้ง ก่อนการรับสมัคร และควรจะมีการมอบหมายให้ชัดเจนว่า หน่วยงานใดจะเป็นผู้ดำเนินการ ประชาสัมพันธ์ ซึ่งควรจะดำเนินการแต่เนิ่น ๆ เพื่อที่ผู้จะมาสมัครจะได้เตรียมตัวได้ทัน ไม่ก่อให้เกิดปัญหาของการมอบหมายภาระงานให้ผู้อื่นทำแทน ซึ่งส่วนนี้จะส่งผลกระทบต่อให้นักศึกษา ต้องออกกลางคัน เนื่องจากยังต้องกลับมารับผิดชอบงานที่ทำอยู่ในระหว่างที่กำลังศึกษาอยู่ใน สถาบันฯ

2.2 การพิจารณาจัดสรรทุนในส่วนของทุนรัฐบาลไทยนั้น ควรจะพิจารณาจัดสรร ลำดับความสำคัญตามลำดับความขาดแคลนบุคลากรในสาขาต่าง ๆ สาขาวิชาใดที่มีความ ขาดแคลนมาก ควรจะได้รับการพิจารณาก่อน และนอกจากนั้น ควรจะพิจารณาด้วยว่าผู้ที่ได้รับทุน ควรจะได้เรียนในสาขาวิชาที่ตรงกับลักษณะงานที่ทำให้มากขึ้น

2.3 กระบวนการจัดสรรทุนในส่วนของสถาบันฯ ปีละ 10 ล้านบาทนั้น ถ้าเป็นไปได้ ควรจะประกาศให้ชัดเจนให้นักศึกษาผู้ได้รับทุน ได้ทราบว่าเป็นทุนภายใต้โครงการความร่วมมือ ระหว่างรัฐบาลไทย และสถาบันฯ และดำเนินการเป็นขั้นตอน เช่นเดียวกับการจัดสรรทุนในส่วน ของรัฐบาลไทย ปีละ 20 ล้านบาท

2.4 ควรพิจารณาเร่งรัดให้นักศึกษาระดับปริญญาเอก สำเร็จการศึกษาภายใน 9 ภาค การศึกษา

เอกสารอ้างอิง

สถาบันเทคโนโลยีแห่งเอเชีย. 2527 *บทบาทของเอไอทีต่อการพัฒนาของประเทศไทย.*

สถาบันเทคโนโลยีแห่งเอเชีย. 2530 *เอไอทีกับการพัฒนาประเทศไทย.*

สถาบันเทคโนโลยีแห่งเอเชีย. 2537 *แนะนำเอไอที.*

สถาบันวิจัยเพื่อการพัฒนาประเทศไทย. 2538 'ยุทธศาสตร์การพัฒนากำลังคนทางเทคโนโลยีเพื่อพัฒนา เศรษฐกิจและความสามารถในการแข่งขัน'. รายงานฉบับสมบูรณ์.

Asian Institute of Technology. 1995 *Student Handbook* 1995/96. Bangkok.

Asian Institute of Technology. *Data for Evaluation of RTG/HRD-AIT Program* 1992-1996.

Development Office. 1995 *Subject Priority for RGT/AIT Scholarship, Asian Institute of Technology.*

Limskul, K., Khantigaroon, T. and Sussangkorn, C. 1992 *Projects of Manpower Usage in Industry and Services. NESDB/UNDP Project.* TDRI. February.

**The Evaluation of the Cooperation
between
the Government of Thailand and the Asian Institute of
Technology as It Pertains to the Thai Government's Budgetary
Support : Phase V (B.E. 2533-2537)**

Somwung Pitiyanuwat ¹ Auyporn Ruengtrakul ¹ Wilaiwan Pitiyanuwat ²
Sopana Takaew ³ Panurut Rattayapas ³ Wantani Thaitiang ³
Tasaneer Sahawat ⁴

ABSTRACT

The purposes of the evaluation of the cooperation between the Government of Thailand and the Asian Institute of Technology as it pertains to the Thai Government's budgetary support: phase V (B.E.2533-2537) are as follows:

- 1. to evaluate the allocation of financial resources.*
- 2. to evaluate the production of graduates under the King's Scholarship Program.*
- 3. to follow up the impact of such cooperation on Thailand.*

The results revealed that the Asian Institute of Technology properly allocated financial resources as planned. Thirty percent of scholarship recipients was Thai and the rest was people in the region. Most of the scholarship that Thai had got were very rare and would have positive impact on Thailand too.

¹ Faculty of Education, Chulalongkorn University

² Planning and Development Division, Chulalongkorn University

³ Office of the National Education Commission

⁴ The Asian Institute of Technology

การประเมินความร่วมมือระหว่างรัฐบาลไทยและสถาบันเทคโนโลยีแห่งเอเชียส่วนที่รัฐบาลไทยให้การสนับสนุนแก่สถาบันเทคโนโลยีแห่งเอเชียตามแผนการอุดหนุน 5 ปี ระยะที่ 5 (พ.ศ. 2533-2537)

สมหวัง พิธิยานุวัฒน์¹ อวยพร เรืองตระกูล¹ วิไลวรรณ พิธิยานุวัฒน์²
โสภณา ตาแก้ว³ ภาณุรัตน์ รัตยาภาส³ วันทนีย์ ไทยเที่ยง³
ทัศนีย์ สหวัฒน์⁴

บทคัดย่อ

การประเมินความร่วมมือระหว่างรัฐบาลไทยและสถาบันเทคโนโลยีแห่งเอเชียในส่วนที่รัฐบาลไทยให้การสนับสนุนแก่สถาบันเทคโนโลยีแห่งเอเชีย ตามแผนการอุดหนุน 5 ปี ระยะที่ 5 (พ.ศ. 2533-2537) มีวัตถุประสงค์เพื่อ (1) ประเมินการใช้จ่ายงบประมาณ (2) ประเมินการผลิตบัณฑิตทุนเล่าเรียนหลวง (3) ติดตามผลกระทบที่มีต่อประเทศไทย

ผลการประเมินปรากฏว่า สถาบันฯ มีการใช้จ่ายงบประมาณเป็นไปตามเจตนารมณ์ที่กำหนด การผลิตบัณฑิตทุนเล่าเรียนหลวงสอดคล้องกับวัตถุประสงค์ของการให้ทุน โดยมีคนไทยได้รับทุนร้อยละ 30 นอกนั้นเป็นชาวต่างชาติที่มาจากภูมิภาคเอเชีย สาขาที่ได้รับทุนเป็นสาขาที่ขาดแคลน และมีผลกระทบทางบวกต่อประเทศไทยในทุก ๆ ด้าน

¹ คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

² ฝ่ายวางแผนและพัฒนา จุฬาลงกรณ์มหาวิทยาลัย

³ สำนักงานคณะกรรมการการศึกษาแห่งชาติ

⁴ สถาบันเทคโนโลยีแห่งเอเชีย

ความเป็นมาของการประเมิน

สืบเนื่องจากสถาบันเทคโนโลยีแห่งเอเชียซึ่งได้จัดตั้งเป็นครั้งแรกใน ปี พ.ศ. 2502 มีฐานะเป็นสถาบันบัณฑิตวิศวกรรม สปอ.(SEATO) เริ่มเปิดรับนักศึกษารุ่นแรกจำนวน 18 คน เป็นนักศึกษาจาก 3 ประเทศ คือ ปากีสถาน ฟิลิปปินส์ และไทย โดยมีทุนสนับสนุนหลักจากประเทศออสเตรเลีย ฝรั่งเศส นิวซีแลนด์ ไทย สหราชอาณาจักรและสหรัฐอเมริกา และทุนอุดหนุนเพิ่มเติมจากประเทศอื่น ๆ ในภูมิภาค คือ ประเทศปากีสถาน และประเทศฟิลิปปินส์

ต่อมาปี พ.ศ. 2509 คณะมนตรีองค์การสนธิสัญญาป้องกันแห่งเอเชียอาคเนย์ร่วมพบปะ ณ กรุงแคนเบอร์รา และมีมติให้ก่อตั้งองค์การอิสระระหว่างประเทศขึ้น เรียกชื่อว่า สถาบันเทคโนโลยีแห่งเอเชีย (Asian Institute of Technology : AIT) ในปี พ.ศ.2510 มีการตราพระราชบัญญัติคุ้มครองการดำเนินงานของสถาบันเทคโนโลยีแห่งเอเชีย และให้สถาบันบัณฑิตวิศวกรรม สปอ. สิ้นสุดลงตั้งนั้น สถาบันเทคโนโลยีแห่งเอเชียจึงเป็นสถาบันอิสระระหว่างชาติ และมีการประกาศกฎบัตรของสถาบันในราชกิจจานุเบกษา เมื่อวันที่ 25 ตุลาคม 2510 โดยรัฐบาลไทยให้สถาบันเทคโนโลยีแห่งเอเชียใช้ที่ดิน 1,100 ไร่ ตั้งอยู่ที่ถนนวิภาวดีรังสิต ในลักษณะเกือบจะเป็นการให้เปล่า ส่วนอาคารและสิ่งปลูกสร้างอื่น ๆ ได้รับเงินอุดหนุนจากรัฐบาลออสเตรเลีย ได้หวัน ญี่ปุ่น นิวซีแลนด์ ประเทศไทย สหราชอาณาจักร สหรัฐอเมริกา และสิงคโปร์

สถาบันเทคโนโลยีแห่งเอเชีย ดำเนินกิจการโดยอาศัยเงินอุดหนุนจากองค์การ มูลนิธิ หน่วยงานต่าง ๆ และรัฐบาลประเทศต่าง ๆ รวมทั้งรัฐบาลไทย โดยมีสำนักงานคณะกรรมการการศึกษาแห่งชาติ เป็นหน่วยงานประสานงานการให้การอุดหนุนแก่สถาบันเทคโนโลยีแห่งเอเชียในนามรัฐบาลไทย

นับจากการก่อตั้งสถาบันมาเป็นเวลากว่า 30 ปี ในประเทศไทย รัฐบาลไทยได้ให้ความอุปถัมภ์และช่วยเหลือแก่สถาบันมาโดยต่อเนื่อง ทั้งนี้เพื่อให้สถาบันได้บรรลุเป้าหมายในการพัฒนาการศึกษาตลอดจนการวิจัยในโครงการพัฒนาต่าง ๆ อันเป็นประโยชน์ต่อประเทศไทยและประเทศกำลังพัฒนาอื่น ๆ โดยมุ่งเน้นในการพัฒนาบุคลากรตลอดจนส่งเสริมความก้าวหน้าทางด้านเทคโนโลยีและการพัฒนา ที่ให้แก่สถาบันการศึกษาของประเทศ รวมทั้งความร่วมมือในโครงการพัฒนาบุคลากรทางด้านวิศวกรรมศาสตร์และเทคโนโลยี และความร่วมมือกับผู้เชี่ยวชาญในคณะรัฐบาลไทย ทราบจนปัจจุบันประเทศไทยเป็นประเทศที่ให้ความช่วยเหลือด้านการเงินแก่สถาบันเทคโนโลยีแห่งเอเชียอย่างต่อเนื่อง รวม 5 แผน โดยเฉพาะแผนระยะที่ 5 ระหว่าง พ.ศ.2533 ถึง 2537 และในระยะเวลาที่ผ่านมา ยังไม่มีการประเมินผลทำให้ความร่วมมือที่รัฐบาลไทยให้แก่สถาบันว่าควรมีทิศทางและรูปแบบการสนับสนุนช่วยเหลือในอนาคตอย่างไร

ดังนั้น เพื่อให้การสนับสนุนช่วยเหลือของรัฐบาลไทยในส่วนของ การอุดหนุนตามแผนปฏิบัติการ 5 ปี ดำเนินไปอย่างมีประสิทธิภาพอันเป็นไปเพื่อประโยชน์ต่อประเทศไทย และประเทศเพื่อนบ้านในภูมิภาคเอเชีย สำนักงานคณะกรรมการการศึกษาแห่งชาติ เห็นสมควรให้มีการประเมินความร่วมมือระหว่างรัฐบาลไทยและสถาบันเทคโนโลยีแห่งเอเชีย ในส่วนที่รัฐบาลไทยให้ความสนับสนุนแก่สถาบันเทคโนโลยีแห่งเอเชีย ตามแผนการอุดหนุน 5 ปี ระยะที่ 5 ระหว่างปี พ.ศ. 2533-2537

วัตถุประสงค์ของการประเมิน

เพื่อติดตามและประเมินผลความร่วมมือที่รัฐบาลไทยได้ให้การสนับสนุนแก่สถาบันเทคโนโลยีแห่งเอเชียตามแผนการอุดหนุน 5 ปี ระยะที่ 5 (พ.ศ. 2533-2537) เกี่ยวกับผลการดำเนินงานดังนี้

- (1) การใช้จ่ายงบประมาณเป็นไปตามเจตนารมณ์ที่กำหนดหรือไม่ เพียงใด มีประสิทธิผลและความคุ้มค่าแก่การดำเนินงานอย่างไร
- (2) ผลการผลิิตบัณฑิตทุนเล่าเรียนหลวง ได้สนองความต้องการของประเทศไทย มากน้อยเพียงใด
- (3) ผลกระทบที่มีต่อประเทศไทยในการที่รัฐบาลได้ให้การสนับสนุนสถาบันเทคโนโลยีแห่งเอเชีย

ขอบเขตของการประเมิน

ในการประเมินครั้งนี้จะติดตามและประเมินผลเฉพาะในส่วนของความร่วมมือที่รัฐบาลไทยให้การสนับสนุนแก่สถาบันเทคโนโลยีแห่งเอเชีย ในช่วงแผนระยะที่ 5 ที่ผ่านมา (พ.ศ. 2533-2537) เท่านั้น โดยจะทำการศึกษาติดตามและประเมินในด้านต่าง ๆ ดังนี้

1. ด้านการดำเนินงานและการใช้จ่ายงบประมาณ
 - 1.1 รายรับรายจ่ายของสถาบันในช่วง 5 ปี (พ.ศ.2533-2537)
 - 1.2 กระบวนการจัดทำแผนและกระบวนการจัดทำงบประมาณ
 - 1.3 การเสนอขอของงบประมาณ
 - 1.4 กระบวนการเบิกจ่ายและการใช้งบประมาณ
 - 1.5 จำนวนอาจารย์ต่อนักศึกษา ค่าใช้จ่ายต่อนักศึกษา และค่าใช้จ่ายต่อทุนการศึกษา
2. ด้านผลการผลิตบัณฑิตทุนเล่าเรียนหลวง
 - 2.1 กระบวนการให้ทุนเล่าเรียนหลวงและทุนภายในของสถาบันเทคโนโลยีแห่งเอเชีย
 - 2.2 จำนวนทุนเล่าเรียนหลวงตามแผนและการปฏิบัติจริง

- 2.3 ประสิทธิภาพในการผลิตบัณฑิตทุนเล่าเรียนหลวง
- 2.4 คุณภาพบัณฑิตทุนเล่าเรียนหลวง
3. ด้านการติดตามผลผู้รับทุนเล่าเรียนหลวง จำนวน 69 คน โดยติดตามผลใน 3 ประเด็น คือ
 - 3.1 ความคิดเห็นเกี่ยวกับทุนเล่าเรียนหลวง
 - 3.2 ผลที่ได้จากการรับทุนเล่าเรียนหลวง
 - 3.3 ผลกระทบของทุนเล่าเรียนหลวง

ทั้งนี้ไม่รวมงบประมาณที่รัฐบาลไทยจัดสรรให้โครงการพัฒนาบุคลากรทางวิศวกรรมศาสตร์และเทคโนโลยีที่สถาบันเทคโนโลยีแห่งเอเชียได้รับจากรัฐบาลไทยในช่วงปี 2535-2537 ปีละ 20 ล้านบาท รวม 60 ล้านบาท

ข้อตกลงเบื้องต้น

1. ข้อมูลและสถิติต่าง ๆ เกี่ยวกับรายรับและรายจ่ายของสถาบัน การจัดสรรทุน ตลอดจนการผลิตบัณฑิตทุนเล่าเรียนหลวง เป็นข้อมูลและสถิติที่ถูกต้องตามสภาพความเป็นจริง
2. ผู้ให้ข้อมูลหรือให้สัมภาษณ์เกี่ยวกับการจัดสรรทุนและการผลิตบัณฑิต ให้ข้อมูลตามประสบการณ์หรือความรู้สึกที่แท้จริง
3. ผู้รับทุนเล่าเรียนหลวงตอบแบบสอบถามตามประสบการณ์และความรู้สึกที่แท้จริง
4. การเก็บข้อมูลในช่วงวัน เวลาที่ต่างกัน ไม่มีผลต่อการให้ข้อมูล
5. นักศึกษาของสถาบันเทคโนโลยีแห่งเอเชียเป็นนักศึกษาเต็มเวลา

คำนิยามเชิงปฏิบัติการ

สถาบัน หมายถึง สถาบันเทคโนโลยีแห่งเอเชีย

ปีงบประมาณของสถาบัน หมายถึง ปีงบประมาณของสถาบันเทคโนโลยีแห่งเอเชีย ซึ่งเริ่มเดือนกรกฎาคมถึงเดือนมิถุนายน

หมวดเงินอุดหนุน หมายถึง เงินอุดหนุนดำเนินการ เงินอุดหนุนสมทบค่าก่อสร้าง และทุนการศึกษา

หมวดรายจ่ายอื่น หมายถึง ภาษีเงินได้ประจำปี ภาษีค้างจ่าย ภาษีอากรรถยนต์ ภาษีอากรของใช้ภายในบ้าน และภาษีอากรสุรา บุหรี่

ค่าใช้จ่ายต่อนักศึกษา =
$$\frac{\text{ค่าใช้จ่ายดำเนินการผลิตบัณฑิต (รายการ 1-7)}}{\text{จำนวนนักศึกษา}}$$

แผนการอุดหนุนระยะที่ 5 หมายถึง ความร่วมมือที่รัฐบาลไทยให้การสนับสนุนแก่สถาบันเทคโนโลยีแห่งเอเชียในช่วง พ.ศ.2533-2537

ทุนเล่าเรียนหลวง หมายถึง ทุนพระราชทานในพระบาทสมเด็จพระเจ้าอยู่หัวที่ให้แก่นักศึกษาในระดับปริญญาโทของสถาบันเทคโนโลยีแห่งเอเชีย

ทุนของสถาบัน หมายถึง ทุนที่มีอยู่ในสถาบันเทคโนโลยีแห่งเอเชียทั้งหมด มีดังต่อไปนี้

1. ทุนภายใน (Internal Scholarships) หมายถึง ทุนที่รัฐบาลหรือหน่วยงานในภาครัฐหรือเอกชนจากประเทศต่าง ๆ รวมทั้งทุนเล่าเรียนหลวงที่มอบให้แก่นักศึกษาของสถาบันเทคโนโลยีแห่งเอเชีย โดยให้สถาบันเทคโนโลยีแห่งเอเชียเป็นผู้ดำเนินการคัดเลือกนักศึกษาที่เก่งที่สุดเข้ารับทุนตามเงื่อนไขต่าง ๆ ที่เจ้าของทุนได้กำหนดเอาไว้

2. ทุนภายนอก (External Scholarships) หมายถึง ทุนที่รัฐบาลหรือหน่วยงานในภาครัฐและเอกชนจากประเทศต่าง ๆ รวมทั้งประเทศไทย (RGT) มอบให้แก่สถาบันเทคโนโลยีแห่งเอเชียโดยทางเจ้าของทุนเป็นผู้จัดส่งคนของตนมาเรียน หรืออาจจะให้ทางสถาบันเทคโนโลยีแห่งเอเชียเป็นผู้ดำเนินการคัดเลือกนักศึกษาที่มีคุณสมบัติตามที่เจ้าของทุนระบุแล้วเจ้าของทุนจะพิจารณาอีกครั้ง

3. ทุนส่วนตัว (Self-paying) หมายถึง ทุนที่นักศึกษาเป็นผู้จัดหาเอง ซึ่งอาจจะได้มาจากหน่วยงานหรือองค์กรใดก็ได้ ทุนบางส่วน (Partial-Scholarship) ก็จัดเป็นทุนประเภทนี้ด้วย ทุนบางส่วนเป็นทุนที่ทางสถาบัน หาให้แก่นักศึกษาที่เข้ามาเรียน 2 ภาคการศึกษา ด้วยทุนส่วนตัวแล้วมีผลการเรียนดี ทางสถาบันจึงหาทุนให้เรียนอีก 3 ภาคการศึกษา

4. ทุนยกเว้นค่าเล่าเรียน (Fee Waiver) หมายถึง ทุนที่สถาบันให้แก่นักศึกษา โดยการยกเว้นไม่เก็บค่าเล่าเรียนสำหรับเจ้าหน้าที่ที่ทำงานอยู่ที่สถาบัน ครอบครัวของอาจารย์ชาวต่างประเทศของสถาบัน รวมทั้งนักศึกษาที่จ่ายเงินค่าเล่าเรียนเองหลังจากที่เรียนไปแล้ว 3 ภาคการศึกษาแล้วมีผลการเรียนดี นักศึกษาจะได้รับยกเว้นค่าเล่าเรียน

ปีการศึกษา หมายถึง ปีการศึกษาของสถาบัน ซึ่งเริ่มตั้งแต่ภาคเรียนเดือนมกราคม เดือนพฤษภาคม จนถึงภาคเรียนเดือนกันยายน รวมแต่ละปีการศึกษามี 3 ภาคการศึกษา

ระเบียบวิธีการประเมิน

การประเมินครั้งนี้ดำเนินการประเมินตามกระบวนการทัศน์ การประเมินเพื่อการจัดการโดยคำนึงถึงความต้องการใช้ผลประเมินเพื่อการตัดสินใจเป็นสำคัญ สำหรับวิธีการประเมินมีรายละเอียดดังนี้

1. การประเมินค่าใช้จ่ายของสถาบันเทคโนโลยีแห่งเอเชียที่มุ่งศึกษาถึงกระบวนการงบประมาณที่รัฐบาลไทยได้ให้การสนับสนุนต่อสถาบันและเสนอแนวทางการดำเนินงานด้านงบประมาณ โดยได้ทำการศึกษา

1.1 รายรับและรายจ่ายของสถาบัน ในช่วงแผน 5 ปี (พ.ศ.2533-2537)

1.2 กระบวนการจัดทำแผนและกระบวนการงบประมาณของงบประมาณที่รัฐบาลไทยให้การสนับสนุน

1.3 การเสนอของบประมาณ การจัดสรรงบประมาณ ที่รัฐบาลไทยให้การสนับสนุน

1.4 กระบวนการเบิกจ่ายและการใช้จ่ายงบประมาณที่รัฐบาลไทยให้การสนับสนุน

1.5 จำนวนอาจารย์ต่อนักศึกษา ค่าใช้จ่ายต่อนักศึกษา ค่าใช้จ่ายต่อทุนการศึกษาของสถาบันโดยเปรียบเทียบกับหน่วยงานอื่น/ทุนประเภทอื่น

1.6 เสนอแนะแนวทางการดำเนินงานในกระบวนการงบประมาณที่รัฐบาลไทยให้การสนับสนุน

2. การประเมินประสิทธิภาพการผลิตบัณฑิตทุนเล่าเรียนหลวงศึกษาในประเด็นต่อไปนี้

2.1 กระบวนการให้ทุนภายในของสถาบันเทคโนโลยีแห่งเอเชีย

2.2 จำนวนทุนเล่าเรียนหลวง

2.3 ประสิทธิภาพการผลิตบัณฑิต

2.4 คุณภาพของบัณฑิต

3. การติดตามผู้รับทุนเล่าเรียนหลวงรุ่นปีการศึกษา 2533-2537 จำนวน 69 คน โดยมุ่งติดตามผลผู้รับทุนเล่าเรียนหลวงในด้านความคิดเห็นเกี่ยวกับทุนเล่าเรียนหลวง ด้านผลที่ได้รับจากทุนเล่าเรียนหลวงและด้านผลกระทบของทุนเล่าเรียนหลวง

วิธีการเก็บรวบรวมข้อมูล

1. การประเมินค่าใช้จ่ายของสถาบัน

เก็บรวบรวมข้อมูลจากเอกสารพระราชบัญญัติงบประมาณรายจ่ายประจำปีงบประมาณ พ.ศ.2533-2537 เอกสารแผนขอรับการสนับสนุนจากรัฐบาลไทย รายงานประจำปีของสถาบัน รายงานการเงิน และเอกสารการเบิกจ่ายของสถาบัน พร้อมกับสัมภาษณ์เพิ่มเติมจากบุคคลที่เกี่ยวข้องของสถาบัน

2. การประเมินประสิทธิภาพการผลิตบัณฑิตทุนเล่าเรียนหลวง

เก็บรวบรวมข้อมูลในด้านการรับสมัคร จำนวนผู้สมัคร จำนวนนักศึกษา จำนวนอาจารย์ จำนวนทุน เกรดเฉลี่ย จำนวนผู้สำเร็จการศึกษา เป็นต้น โดยคัดลอกข้อมูลจากฝ่ายรับนักศึกษา

ฝ่ายทะเบียนนักศึกษา รวมทั้งเจ้าหน้าที่ฝ่ายทุนการศึกษา สำหรับข้อมูลในเรื่องกระบวนการจัดสรร ทุน เกณฑ์การคัดเลือกทุน ปัญหา อุปสรรคในการจัดสรรทุน เป็นต้น ได้จากการวิเคราะห์เอกสาร และสัมภาษณ์ผู้ที่เกี่ยวข้องกับการให้ทุน คือคณบดีแต่ละคณะ และเจ้าหน้าที่ผู้ดูแลทุน ส่วนสาเหตุ ของการออกกลางคันได้จากการสัมภาษณ์ผู้ที่ออกกลางคัน

ในกรณีข้อมูลของทุนเล่าเรียนหลวงได้เก็บรวบรวมมา 100 % สำหรับทุนภายในอื่น ๆ ที่ นำมาเปรียบเทียบได้เก็บข้อมูลในแต่ละปี (ปี 2533-2537) ร้อยละ 65.42,81.19,72.49,33.32 (ไม่ ได้เก็บข้อมูล เดือนพฤษภาคม พ.ศ.2536 เนื่องจากไม่มีทุนเล่าเรียนหลวงในภาคการศึกษานี้) และ 57.00 โดยเฉลี่ยคิดเป็นร้อยละ 69.28

3. การติดตามผู้รับทุนเล่าเรียนหลวง

ส่งแบบสอบถามไปทางไปรษณีย์ให้ผู้รับทุนเล่าเรียนหลวงซึ่งเริ่มรับทุนเข้าเรียนตั้งแต่ปีการ ศึกษา 2533-2537 จำนวนรวม 69 คน (ไม่รวมที่ลาออกระหว่างเรียน 1 คน) ซึ่งในจำนวน 69 คน มีผู้ที่อยู่ระหว่างการรับทุนการศึกษา 12 คน ได้รับแบบสอบถามคืนจำนวน 52 ฉบับ คิดเป็นร้อยละ 75.36 ของแบบสอบถามที่ส่ง

วิธีการวิเคราะห์ข้อมูล

1. การประเมินค่าใช้จ่ายของสถาบัน

1.1 รายรับและรายจ่ายของสถาบันฯ วิเคราะห์ประเภทและแหล่งที่ให้การสนับสนุน จำนวนงบประมาณของสถาบัน และประเภทรายจ่ายและจำนวนรายจ่ายของสถาบันจากเอกสาร รายงานประจำปีและรายงานการเงินของสถาบันพร้อมทั้งสัมภาษณ์เพิ่มเติม

1.2 กระบวนการจัดทำแผนและกระบวนการงบประมาณ วิเคราะห์กระบวนการจาก เอกสารแผนการสนับสนุนงบประมาณ ขั้นตอนการทำแผน ขั้นตอนการจัดทำงบประมาณรายปีและ ขั้นตอนการเบิกจ่าย

1.3 การเสนอขอของงบประมาณและจัดสรรงบประมาณ วิเคราะห์เปรียบเทียบงบ งบประมาณและรายการที่กำหนดไว้ในแผนที่เสนอขอรายปี ที่ได้รับจัดสรรจากรัฐบาลไทย จากข้อมูล เอกสารแผนขอรับการสนับสนุนจากรัฐบาลไทย เอกสารพระราชบัญญัติงบประมาณรายจ่ายประจำ ปี

1.4 กระบวนการเบิกจ่ายและการใช้จ่ายวิเคราะห์งบประมาณและรายการที่ขอเบิกจ่าย โดยเปรียบเทียบกับงบประมาณและรายการที่ได้รับจัดสรร ระยะเวลาของการเบิกจ่ายจากเอกสาร ของสถาบันเทคโนโลยีแห่งเอเชียเกี่ยวกับการเบิกจ่ายเงินที่รับการอุดหนุนจากรัฐบาลไทย

1.5 จำนวนอาจารย์ต่อนักศึกษา ค่าใช้จ่ายต่อหัว และค่าใช้จ่ายต่อทุนของสถาบัน โดยวิเคราะห์หาอัตราส่วนอาจารย์ต่อนักศึกษาของสถาบันเปรียบเทียบกับเกณฑ์มาตรฐานอาจารย์ต่อนักศึกษาในสถาบัน

2. การประเมินประสิทธิภาพการผลิตบัณฑิตทุนเล่าเรียนหลวงเนื่องจากข้อมูลระดับผลการเรียนในระดับปริญญาตรีของนักศึกษาที่มาขอทุนนั้น มาจากหลายประเทศ และหลายมหาวิทยาลัย ซึ่งมีระบบการให้ระดับผลการเรียนต่างกัน ซึ่งมีทั้งระบบเต็ม 4,5,6 และเต็ม 10 แต้ม บางมหาวิทยาลัยให้ระบบเป็นเปอร์เซ็นต์บางมหาวิทยาลัยให้เป็น A,B, Pass และบางมหาวิทยาลัยให้เป็นระดับที่ของห้อง ดังนั้นทางผู้ประเมินจึงได้เปลี่ยนระบบการให้ผลการเรียนให้เป็นฐานเดียวกันโดยเปลี่ยนในส่วนของระบบเต็ม 5,6 และเต็ม 10 เปลี่ยนให้เป็นเต็ม 4 แต้ม ส่วนผลการเรียนที่เป็นตัวอักษรได้เปลี่ยนเป็นแต้มเต็ม 4 ดังนี้

$$A+=4.00:A=3.75:B+=3.25$$

$$B =3.00:C+=2.25:C=2.00$$

สำหรับเกณฑ์ในการประเมินผลการเรียนครั้งนี้ ได้ตั้งในระบบเต็ม 4 แต้ม ดังนี้

4.00	ดีเยี่ยม
3.50-3.99	ดีมาก
3.00-3.49	ดี
2.50-2.99	พอใช้

ส่วนผู้ที่ได้รับอยู่ในอันดับแรกของห้อง ถือว่าอยู่ในเกณฑ์ที่ดีเยี่ยม 2-5 อันดับแรกของห้องถือว่าอยู่ในเกณฑ์ดีมาก

ในส่วนของกาวิเคราะห์ข้อมูลนั้น ได้ดำเนินการวิเคราะห์ข้อมูลหาค่าสถิติด้วยโปรแกรมสำเร็จรูป EXCEL, SPSS PC⁺ และกาวิเคราะห์เนื้อหา

3. การติดตามผู้รับทุนเล่าเรียนหลวง คำนวณหาค่าสถิติพื้นฐาน โดยใช้โปรแกรม SPSS PC⁺ และ การวิเคราะห์เนื้อหา (Content Analysis) ในส่วนที่เป็นคำถามปลายเปิดและข้อความ

ผลการประเมิน

1. ผลการประเมินค่าใช้จ่ายของสถาบัน แสดงในตารางที่ 1-6

1.1 ในปี 2533-2537 สถาบันได้รับงบประมาณสนับสนุนรวม 2,888 ล้านบาท โดยได้รับการสนับสนุนจากรัฐบาลประเทศต่าง ๆ คิดเป็นร้อยละ 74 จากองค์กรระดับนานาชาติร้อยละ 11 จากองค์กรรัฐระดับประเทศร้อยละ 8 และจากองค์กรธุรกิจเอกชน มูลนิธิ คิดเป็นร้อยละ 8 ประเทศที่ให้การสนับสนุนมาก 5 ประเทศ คือ ไทย สหรัฐอเมริกา ญี่ปุ่น เยอรมัน และแคนาดา

1.2 รายรับและรายจ่ายในการบริหารงานของสถาบันใกล้เคียงกันในแต่ละปี ในช่วง 5 ปี สถาบันมีรายได้เหนือรายจ่ายประมาณ 52 ล้านบาท รายรับส่วนใหญ่มาจากค่าเทอมและค่าบำรุงการศึกษา รายจ่ายส่วนใหญ่เป็นเงินเดือนค่าจ้างอาจารย์และพนักงาน

1.3 การจัดทำแผนของสถาบันมีลักษณะเป็นการจัดทำแผนจากล่างสู่บน คือ สถาบันเป็นผู้นำเสนอสำนักงานคณะกรรมการการศึกษาแห่งชาติตามนโยบายของสถาบัน แผนที่สถาบันเสนอเป็นแผนเสนองบประมาณ 5 ปี มากกว่าเป็นการจัดทำแผนพัฒนาหรือแผนดำเนินงาน ในการจัดทำค่าของงบประมาณรายจ่ายประจำปีของสถาบัน จะมีขั้นตอนและวิธีดำเนินการที่กำหนด โดยกรอบหรือกระบวนการจัดทำงบประมาณแผ่นดิน ทั้งนี้เป็นที่น่าสังเกตว่าคำชี้แจงประกอบการขอและแบบฟอร์มทางสถาบันมีความสะดวกคล่องตัวกว่าของหน่วยราชการหรือรัฐวิสาหกิจอื่น สถาบันเน้นกระบวนการเบิกจ่าย ส่วนกระบวนการติดตามผลการใช้จ่ายและรายงานผลยังไม่ปรากฏ ในโครงสร้างการบริหาร การเบิกจ่ายสถาบันมีความคล่องตัวมากในหมวดเงินอุดหนุน ส่วนในหมวดรายจ่ายอื่นยังต้องเบิกจ่ายเป็นรายการตามหลักฐานการใช้จ่าย

1.4 สถาบันตั้งงบประมาณในหมวดเงินอุดหนุนสูงกว่าแผนและที่รัฐบาลจัดสรรให้ (96.1 ล้านบาท) ซึ่งสูงกว่าแผนและต่ำกว่าที่สถาบันเสนอขอมาเล็กน้อย กล่าวคือ ได้รับการจัดสรรในหมวดอุดหนุนร้อยละ 94 ของงบประมาณที่เสนอขอ ส่วนรายจ่ายอื่น สถาบันได้รับจัดสรร 231.2 ล้านบาท คิดเป็นร้อยละ 108 ของงบประมาณที่เสนอขอ แต่ก็ยังเป็นงบประมาณที่ต่ำกว่าแผน

1.5 ในระยะ 5 ปี สถาบันได้เบิกจ่ายงบประมาณเต็ม 78 ทู่น คิดเป็นเงิน 38.285 ล้านบาท ซึ่งเป็นเงินอุดหนุนทั่วไป จึงเบิกจ่ายได้โดยไม่ต้องมีหลักฐานการเบิกจ่าย สถาบันได้นำเงินนี้ไปจัดสรรทุนให้นักศึกษาจำนวน 70 ทู่น จำนวนทุนที่จัดสรรขาดไป 8 ทู่น สถาบันจะจัดสรรให้ในปีต่อไป เป็นที่น่าสังเกตว่าการจัดสรรทุนเล่าเรียนหลวงของสถาบันในแต่ละปี บางปีจัดสรรมากกว่าจำนวนทุนที่ได้รับจากรัฐบาลไทย แต่บางปีจัดสรรน้อยกว่าทุนที่ได้รับ

1.6 โดยเฉลี่ย 5 ปีสถาบันมีค่าใช้จ่ายงบดำเนินการต่อนักศึกษาใกล้เคียงกันอยู่ระหว่าง 441,000 บาท/คน/ปี ถึง 495,000 บาท/คน/ปี หรือโดยเฉลี่ยเท่ากับ 475,000 บาท/คน/ปี และถ้าไม่รวมงบที่ใช้จ่ายเป็นค่าจ้างอาจารย์และพนักงาน ค่าใช้จ่ายต่อนักศึกษาเป็น 241,000 บาท/คน/ปี หรืออยู่ระหว่าง 205,000 บาท/คน/ปี ถึง 257,000 บาท/คน/ปี เมื่อเทียบค่าใช้จ่ายที่สถาบันลงทุนเพื่อนักศึกษา คนละ 475,000 บาทต่อปี กับทุนนักศึกษาที่สถาบันได้รับ 1 ปี (3 เทอม) เป็นเงิน 289,170 บาทในปี 2533 และ 337,365 บาทในปี 2534 แสดงว่าทุนการศึกษาที่สถาบันได้รับการสนับสนุนจากผู้สมัครไม่เพียงพอต่อการใช้จ่ายในการผลิตบัณฑิต กล่าวคือ ค่าใช้จ่ายต่อผู้สำเร็จการศึกษาส่วนที่สถาบันรับผิดชอบประมาณ 792,000 บาท ต่อผู้สำเร็จการศึกษา ทั้งนี้ยังไม่รวมค่าใช้จ่ายที่จ่ายให้ผู้รับทุนเป็นรายเดือนอีก 84,000 บาทต่อผู้สำเร็จการศึกษา

ตารางที่ 1 อัตราส่วนงบประมาณที่ให้การสนับสนุน สถาบัน ปี 2533-2537 จำแนกตามประเภทหน่วยงานที่ให้การสนับสนุน

ประเภทหน่วยงาน	2533	2534	2535	2536	2537	เฉลี่ย 5 ปี	
						ล้านบาท	%
รัฐบาล	75.72	78.15	75.25	75.19	65.18	2,124.27	73.56
- ประเทศไทย	9.87	8.80	10.55	8.04	9.92	271.82	9.41
- ประเทศสหรัฐอเมริกา	4.49	7.77	7.52	3.95	1.70	143.93	4.98
- ประเทศญี่ปุ่น	10.61	9.05	8.96	10.51	10.89	289.15	10.01
- ประเทศเยอรมัน	7.69	10.09	7.13	9.15	4.55	220.50	7.64
- ประเทศแคนาดา	12.47	7.46	2.25	5.33	3.76	167.57	5.80
- ประเทศอื่นๆ	30.59	34.98	38.84	38.21	34.36	1,031.31	35.71
องค์กรระดับนานาชาติ	10.56	9.33	8.80	8.18	16.86	313.43	10.85
องค์กรภาครัฐระดับประเทศ	7.44	5.89	5.28	8.40	11.69	227.46	7.88
องค์กรธุรกิจ เอกชน มูลนิธิ	6.28	6.63	10.67	8.23	6.27	225.51	7.71
รวม	100.00	100.00	100.00	100.00	100.00		100.00
งบประมาณสนับสนุน (ล้านบาท)	436.848	538.146	607.493	650.058	655.124	2,887.669	

ตารางที่ 2 เปรียบเทียบงบประมาณที่ขอรับอุดหนุนจากรัฐบาลไทย ที่ขอในแผนและที่ได้รับอนุมัติแผน

ประเภทรายจ่าย	2533			2534			2535			2536			2537			รวม		
	ขอแผน	อนุมัติ แผน	%	ขอแผน	อนุมัติ แผน	%	ขอแผน	อนุมัติ แผน	%	ขอแผน	อนุมัติ แผน	%	ขอแผน	อนุมัติ แผน	%	ขอแผน	อนุมัติ แผน	%
1. หมวดเงินอุดหนุน	18.020	18.020	100.00	18.020	18.020	100.00	20.950	20.020	95.56	20.950	20.020	95.56	24.360	20.020	82.18	102.300	96.100	93.94
1.1 งบอุดหนุนการดำเนินการ	5.000	5.000		5.000	5.000		6.000	6.000		6.000	6.000		7.000	6.000		29.000	28.000	
1.2 งบสมทบค่าก่อสร้าง	6.000	6.000		6.000	6.000		7.000	7.000		7.000	7.000		8.000	7.000		34.000	33.000	
1.3 ทุนเล่าเรียนหลวง	7.020	7.020		7.020	7.020		7.950	7.020		7.950	7.020		9.360	7.020		39.300	35.100	
2. หมวดรายจ่ายอื่น	35.000	35.000	100.00	40.000	40.000	100.00	45.000	45.930	102.07	50.000	50.930	102.07	55.000	59.340	101.86	225.000	231.200	107.89
2.1 เงินชดเชยภาษี	35.000	35.000		40.000	40.000		45.000	45.930		50.000	50.930		55.000	59.340		225.000	231.200	
รวม	53.020	53.020	100.00	58.020	58.020	100.00	65.950	65.950	100.00	70.950	70.950	100.00	79.360	79.360	100.00	327.300	327.300	100.00

ตารางที่ 3 เปรียบเทียบงบประมาณที่สถาบันได้รับอุดหนุนจากรัฐบาลไทยตามที่ขอ ได้รับ และเบิกจ่าย

ประเภทรายจ่าย	ขอ	ได้รับ	เบิกจ่ายใน ปีงบประมาณ	เบิกจ่าย ปีถัดไป	รวม เบิกจ่าย	ร้อยละของเบิกจ่าย ต่อได้รับ
1. หมวดเงินอุดหนุน	103.294	97.285	110.285	1.000	111.285	114.39
1.1 งบอุดหนุนการดำเนินการ	33.000	27.000	31.000		31.000	114.81
1.2 งบสมทบค่าก่อสร้าง	32.000	32.000	36.000		36.000	112.50
1.3 ทุนเล่าเรียนหลวง	38.294	38.285	38.285		38.285	100.00
1.4 ระบบข้อมูลชนบทฯ	0.000	0.000	5.000	1.000	6.000	
2. หมวดรายจ่ายอื่น	153.450	123.753	80.169	8.018	88.187	71.26
2.1 ภาษีค้างจ่าย	30.600	22.316	9.857	5.442	15.299	68.56
2.2 ภาษีเงินได้ประจำปี	71.400	63.807	57.000	1.840	58.840	92.22
2.3 ภาษีอากรรถยนต์	33.250	28.550	9.926	0.000	9.926	34.77
2.4 ภาษีอากรของใช้ในบ้าน	5.000	4.400	0.028	0.000	0.028	0.64
2.5 ภาษีอากรสุรา	13.200	4.680	3.358	0.736	4.094	87.48
รวม	256.744	221.038	190.454	9.018	199.472	90.24

ตารางที่ 4 จำนวนทุนเล่าเรียนหลวงในปี 2533-2537

		2533	2534	2535	2536	2537	รวม	%แผน	%ได้รับ
ขอแผน	จำนวนทุน	15	15	17	17	20	84		
	งบประมาณ	7,020,000	7,020,000	7,956,000	7,956,000	9,360,000	39,312,000		
อนุมัติแผน	จำนวนทุน	15	15	15	15	15	75	89.29	
	งบประมาณ	7,020,000	7,020,000	7,020,000	7,020,000	7,020,000	35,100,000	89.29	
ขอรายปี	จำนวนทุน	15	15	15	15	18	78	92.86	
	งบประมาณ	6,885,000	6,885,000	6,885,000	8,000,000	9,639,000	38,294,000	97.41	
ได้รับ(อนุมัติรายปี)	จำนวนทุน	15	15	15	15	18	78	92.86	
	งบประมาณ	6,885,000	6,885,000	6,885,000	8,000,000	9,630,000	38,285,000	97.39	
สถาบันจัดสรร	จำนวนทุน	15	17	9	11	18	70	83.33	89.74
	งบประมาณ	7,400,115	8,963,397	4,327,392	5,745,616	4,873,339	31,309,859	79.64	81.78
ได้รับ-จัดสรร							6,975,141		

ตารางที่ 5 ค่าใช้จ่ายต่อนักศึกษาของสถาบัน

	ประเภทของรายจ่าย	2533		2534		2535		2536		2537		รวม	
		ล้านบาท	%	ล้านบาท	%	ล้านบาท	%	ล้านบาท	%	ล้านบาท	%	ล้านบาท	%
1	เงินเดือนและค่าจ้าง	189.044	45.75	206.382	44.72	221.321	38.57	234.131	39.94	231.695	38.22	1,082.573	40.99
	- เงินสด	116.850		130.004		148.646		165.161		161.964		722.625	
	- การช่วยเหลือซึ่งประเมินเป็นตัวเงิน	72.194		76.378		72.675		68.970		69.731		359.948	
2	ค่าใช้จ่ายเกี่ยวกับการศึกษา	75.915	18.37	96.647	20.94	113.146	19.72	130.718	22.30	142.153	23.45	558.579	21.15
3	ค่าใช้จ่ายเกี่ยวกับนักศึกษา	6.013	1.46	7.065	1.53	7.344	1.28	7.739	1.32	7.682	1.27	35.843	1.36
4.	ค่าใช้จ่ายในการดำเนินงาน	2.521	0.61	14.828	3.21	3.850	0.67	4.252	0.73	4.491	0.74	29.942	1.13
5.	ค่าใช้จ่ายในการซื้อสินทรัพย์ถาวร	31.003	7.50	11.622	2.52	10.118	1.76	16.935	2.89	12.432	2.05	82.110	3.11
6.	ค่าใช้จ่ายเกี่ยวกับมหาวิทยาลัย	33.679	8.15	40.831	8.85	55.701	9.71	49.890	8.51	56.693	9.35	236.794	8.97
7.	ค่าใช้จ่ายเกี่ยวกับอาคาร	14.496	3.51	30.723	6.66	49.366	8.60	40.059	6.83	35.514	5.86	170.158	6.44
8.	ค่าใช้จ่ายโครงการวิจัยและสัมมนา	3.404	0.82	1.910	0.41	9.641	1.68	5.916	1.01	10.844	1.79	31.715	1.20
9.	ค่าใช้จ่ายด้านการบริการสำหรับวิสาหนักจยอย	46.661	11.29	36.670	7.95	86.311	15.04	63.797	10.88	66.665	11.00	300.104	11.36
10.	ขาดทุนจากอัตราแลกเปลี่ยนเงินตราต่างประเทศ	6.908	1.67	7.793	1.69	2.578	0.45	5.316	0.91	1.215	0.20	23.810	0.90
11.	ค่าใช้จ่ายอื่น	3.563	0.86	6.990	1.51	14.446	2.52	27.428	4.68	36.870	6.08	89.297	3.38
12.	รวมรายจ่ายทั้งสิ้น	413.207	100.00	461.461	100.00	573.822	100.00	586.181	100.00	606.254	100.00	2,640.925	100.00
	ประเภทไม่จำกัดขอบเขต	178.218	43.13	213.607	46.29	235.920	41.11	244.268	41.67	257.579	42.49	1,129.592	42.77
	ประเภทจำกัดขอบเขต	234.989	56.87	247.854	53.71	337.902	58.89	341.913	58.33	348.675	57.51	1,511.333	57.23
	รวม 1-7	352.671	85.350	408.098	88.436	460.846	80.312	483.724	82.521	490.660	80.933	2,195.999	83.153
	นักศึกษา	800		880		931		979		1,038		4,628	
	ค่าใช้จ่ายต่อหัว (1-7)	0.441		0.464		0.495		0.494		0.473		0.475	
	รวม 2-7	163.627	39.599	201.716	43.712	239.525	41.742	249.593	42.580	258.965	42.716	1,113.426	42.160
	ค่าใช้จ่ายต่อหัว (2-7)	0.205		0.229		0.257		0.255		0.249		0.241	

ตารางที่ 6 ค่าใช้จ่ายต่อทุนการศึกษาระดับปริญญาโท

2533	อเมริกา	อังกฤษ	ออสเตรเลีย	แคนาดา	นิวซีแลนด์	เยอรมัน	ญี่ปุ่น	ฝรั่งเศส	นอร์เวย์	สถาบัน AIT
1. ค่าใช้จ่ายประจำเดือนตลอดปี	332,640	267,456	279,360	332,640	168,300	258,960	499,422	293,760	377,400	50,400
2. ค่าเล่าเรียน	340,200	346,260	349,200	340,200	495,000	-	113,850	19,200	3,400	237,150
3. ค่าหนังสือและอุปกรณ์การศึกษา	18,900	15,920	11,640	18,900	10,230	12,616	19,481	16,128	22,355	20,043
4. ค่าธรรมเนียมเพื่อการศึกษา	-	-	10,670	-	-	-	-	-	-	-
5. ค่าเบี้ยประกันสุขภาพ	15,120	-	5,238	15,120	5,280	16,434	5,060	24,000	-	-
6. ค่าใช้จ่ายเบ็ดเตล็ด	3,780	11,940	2,134	3,780	-	9,960	13,536	-	-	-
7. เงินชดเชยค่าใช้จ่ายในการดูแลนักเรียน	4,032	4,776	3,104	4,032	2,970	4,980	7,109	4,800	4,144	-
รวมค่าใช้จ่ายต่อคนต่อปี	714,672	646,352	661,346	714,672	681,780	302,950	658,458	357,888	407,299	307,593

2. ผลการประเมินการผลิตบัณฑิตทุนเล่าเรียนหลวง

2.1 ทุนเล่าเรียนหลวงเป็นทุนเพื่อพัฒนากำลังคนระดับสูง เพื่อพัฒนาประเทศในภูมิภาคเอเชีย โดยเน้นที่ประเทศไทยและประเทศที่กำลังพัฒนา โดยคาดหวังว่าผู้รับทุนเล่าเรียนหลวงจะเป็นผู้มีผลการเรียนดีเยี่ยม เป็นการให้ทุนโดยไม่มีเงื่อนไขและไม่ได้กำหนดสาขา สถาบัน จัดทุนเล่าเรียนหลวงเป็นทุนภายในของสถาบัน จึงทำให้ผู้รับทุนเล่าเรียนหลวงไม่ทราบล่วงหน้า จะทราบว่าได้ทุนเล่าเรียนหลวงหลังจากมาเรียนแล้วหลายสัปดาห์ สถาบันดำเนินการคัดเลือกนักศึกษาเข้ารับทุนเล่าเรียนหลวงเช่นเดียวกับทุนภายในอื่น ๆ ของสถาบัน โดยจะจัดสรรทุนให้แก่ทุนภายในที่มีเงื่อนไขก่อน ที่เหลือสุดท้ายจึงจัดให้แก่ทุนภายในที่ไม่มีเงื่อนไข คือทุนเล่าเรียนหลวง การพิจารณาผู้รับทุนภายในไม่มีการสอบแข่งขัน จะใช้การพิจารณาคัดสรรจากประสบการณ์ของอาจารย์ในแต่ละสาขา โดยอาศัยแฟ้มเฉลี่ยสะสมในระดับปริญญาตรี พื้นความรู้ ประสบการณ์ทำงาน จุดหมายแนะนำ และผลการสอบภาษาอังกฤษ เป็นเกณฑ์พิจารณา โดยทั่วไปผู้ที่มีผลการเรียนไม่ต่ำกว่า 3.00 ก็จะได้ทุนเกือบทุกคน

2.2 ในตลอดแผนอุดหนุนระยะที่ 5 สถาบันจัดสรรทุนเล่าเรียนหลวงไป 70 ทุน จากทุนทั้งหมด 78 ทุน โดยจัดสรรทุนให้นักศึกษาชาวไทยร้อยละ 30 และนักศึกษาชาวต่างประเทศร้อยละ 70 ซึ่งสอดคล้องกับวัตถุประสงค์ของทุน คือให้ทุนแก่ประเทศในภูมิภาคเอเชียโดยเน้นประเทศไทย ในเชิงคุณสมบัติผู้รับทุนนั้นปรากฏว่าผู้รับทุนเล่าเรียนหลวงเทียบร้อยละ 28 เท่านั้นที่มีผลการเรียนในระดับปริญญาตรีในระดับดีมาก หรือดีเยี่ยม ในเชิงเปรียบเทียบพบว่า นักศึกษาทุนเล่าเรียนหลวงมีผลการเรียนในระดับปริญญาตรีอยู่ในระดับรองลงมาจากนักศึกษาทุนภายในอื่น ๆ ทั้งนี้เพราะกระบวนการจัดสรรทุนภายในของสถาบันจะจัดสรรให้ทุนที่มีเงื่อนไขก่อน ในกรณีทุนเล่าเรียนหลวงซึ่งไม่มีเงื่อนไขสถาบันจึงจะจัดให้นักศึกษาที่เหลือรับทุนเล่าเรียนหลวง (แสดงในแผนภาพ)

2.3 ประสิทธิภาพในการผลิตบัณฑิตทุนเล่าเรียนหลวงอยู่ในเกณฑ์ดีมาก นักศึกษาทุนเล่าเรียนหลวงร้อยละ 98 สำเร็จการศึกษา ซึ่งนับว่าเป็นอัตราร้อยละของผู้สำเร็จการศึกษาสูงกว่าผู้รับทุนภายในอื่น ๆ ของสถาบัน ทุนโครงการ พสวท. โครงการพัฒนานุเคราะห์ และสาขาการจัดการของสถาบันบัณฑิตธุรกิจศึนทร์แห่งจุฬาลงกรณ์มหาวิทยาลัย บัณฑิตทุนเล่าเรียนหลวงส่วนใหญ่มีผลสัมฤทธิ์ทางการเรียนอยู่ในระดับดี บัณฑิตทุนเล่าเรียนหลวงที่มีผลสัมฤทธิ์ทางการเรียนระดับดีมากหรือดีเยี่ยมมีน้อยกว่าจำนวนบัณฑิตทุนภายในอื่น ๆ ของสถาบัน

3.4 ผู้รับทุนเล่าเรียนหลวงให้เหตุผลที่เลือกมาเรียนที่สถาบันเทคโนโลยีแห่งเอเชีย เพราะชื่อเสียงของสถาบัน การที่สถาบันมีทุนให้ และคุณภาพของอาจารย์ ผู้รับทุนส่วนใหญ่เห็นว่า สถาบันมีหลักสูตรและการเรียนการสอนค่อนข้างดีมาก บัณฑิตทุนเล่าเรียนหลวงสามารถนำความรู้ และประสบการณ์ไปใช้ประโยชน์ในด้านวิชาการ ด้านความรู้ความสามารถเป็นนักวิจัย การปรับตัว ความสามารถเท่ากับการเปลี่ยนแปลงทางเทคโนโลยี และด้านรายได้ในระดับมาก ส่วนด้านการใช้ ภาษาอยู่ในระดับค่อนข้างมาก

3.5 จากการรายงานของผู้รับทุนแสดงว่าทุนเล่าเรียนหลวงมีผลกระทบทางบวกต่อผู้รับทุนอย่างเด่นชัด กล่าวคือ การได้รับทุนเล่าเรียนหลวงทำให้เกิดความเปลี่ยนแปลงด้านต่าง ๆ ทั้ง ด้านตำแหน่งหน้าที่การงาน สถานะทางสังคม สถานะทางเศรษฐกิจและคุณภาพชีวิต ล้วนเปลี่ยนแปลงไปในทิศทางที่ดีขึ้น อีกทั้งส่งผลกระทบในทางบวกต่อประเทศไทย ก่อให้เกิดความร่วมมือและเพิ่มความสัมพันธ์ระหว่างประเทศในเอเชีย ผู้รับทุนจากต่างประเทศจะรู้สึกเป็นหนี้บุญคุณ ประเทศไทย ผู้รับทุนชาวไทยเกิดความภาคภูมิใจ เพิ่มพูนความจงรักภักดีต่อพระบาทสมเด็จพระเจ้าอยู่หัว ตลอดจนเกิดจิตสำนึกผลักดันในการพัฒนาสังคมประเทศไทย บัณฑิตทุนเล่าเรียนหลวงมี ผลงานวิชาการที่เป็นประโยชน์ต่อประเทศชาติ มีการทำวิจัยและนำวิทยานิพนธ์ไปประยุกต์ใช้ใน งานด้านต่าง ๆ รวมทั้งได้สรรสร้างผลงานวิชาการอันเป็นประโยชน์ต่อการพัฒนาประเทศไทย และ ประเทศในภูมิภาคเอเชียตามปณิธานของทุนเล่าเรียนหลวง (แสดงในตารางที่ 7-9)

ตารางที่ 7 ความคิดเห็นเรื่องความรู้และประสบการณ์ที่ได้นำไปใช้ประโยชน์

ด้าน	ผู้รับทุนชาวไทย			ผู้รับทุนชาวต่างประเทศ			รวม		
	n	\bar{X}	S.D	n	\bar{X}	S.D	n	\bar{X}	S.D
วิชาการ	17	4.06	0.75	22	4.05	0.79	39	4.05	0.76
ความรู้และความสามารถ									
เป็นนักวิจัย	17	3.59	1.18	22	4.18	0.96	39	3.92	1.09
ความสามารถทางภาษา	17	3.47	1.28	23	3.30	1.11	40	3.38	1.17
การปรับตัว	17	4.12	0.86	23	3.96	1.11	40	4.03	1.01
ความสามารถเท่าทันการ									
เปลี่ยนแปลงทางเทคโนโลยี	17	4.00	0.94	22	3.95	0.79	39	3.97	0.85
ด้านเงินเดือน/รายได้	16	3.69	1.25	21	3.24	0.81	37	3.54	1.02

คะแนน : 5 - มากที่สุด 4 - มาก 3 - ค่อนข้างมาก 2 - ปานกลาง 1 - น้อย

ตารางที่ 8 หน่วยงานที่ผู้รับทุนเล่าเรียนหลวงที่จบการศึกษาแล้วไปทำงาน

หน่วยงานที่ทำงาน	ผู้รับทุนชาวไทย	ผู้รับทุนชาวต่างประเทศ	รวม
รัฐ	2 (11.76)	7 (30.43)	9 (22.50)
เอกชน	11 (64.71)	13 (56.52)	24 (60.00)
รัฐวิสาหกิจ	1 (5.88)	1 (4.35)	2 (5.00)
ยังไม่ได้ทำงาน	3 (17.65)	2 (8.70)	5 (12.50)
รวม	17 (100.00)	23 (100.00)	40(100.00)

ตารางที่ 9 การเปลี่ยนแปลงด้านต่างๆ

การเปลี่ยนแปลง	ผู้รับทุนชาวไทย (n = 17)			ผู้รับทุนชาวต่างประเทศ (n = 24)			รวม (n = 41*)		
	ดีขึ้น	เหมือนเดิม	ไม่ตอบ	ดีขึ้น	เหมือนเดิม	ไม่ตอบ	ดีขึ้น	เหมือนเดิม	ไม่ตอบ
ตำแหน่งหน้าที่การงาน	11 (64.71)	1 (5.88)	5	14 (58.33)	4 (16.67)	6	25 (60.97)	5 (12.20)	11
สถานะทางสังคม	9 (52.94)	2 (11.76)	6	14 (58.33)	3 (12.50)	7	23 (56.10)		13 (12.20)
สถานะทางเศรษฐกิจ	9 (52.94)	2 (11.76)	6	15 (62.50)	1 (4.17)	8	24 (58.54)	3 (7.32)	14
คุณภาพชีวิต	13 (76.47)	1 (5.88)	3	17 (70.83)	2 (8.33)	5	30 (73.17)	3 (7.32)	8

*ผู้จบการศึกษาแล้ว 57 คน ตอบแบบสอบถาม 41 คน

การที่พระบาทสมเด็จพระเจ้าอยู่หัว พระราชทานทุนเล่าเรียนหลวง ให้กับนักศึกษาไทย และนักศึกษาในภูมิภาคเอเซียนั้น มีผลต่อประเทศไทยในทางที่ดีหลายประการ เช่น ทำให้ประเทศไทยเป็นที่รู้จักมากขึ้น ภาพพจน์ประเทศไทยดีขึ้น เป็นการสร้างและพัฒนาศักยภาพของประเทศไทย ให้มีความสามารถสูงเพิ่มขึ้น ผู้รับทุนเล่าเรียนหลวงเมื่อจบการศึกษาแล้วส่วนใหญ่ได้มีโอกาสทำงานในประเทศไทยระยะหนึ่ง ซึ่งมีผลทำให้เกิดการพัฒนาขึ้นและแก้ปัญหาการขาดแคลนบุคลากรด้านวิทยาศาสตร์และเทคโนโลยีได้บ้าง การให้ทุนแก่นักศึกษาต่างประเทศก็มีผลต่อประเทศไทยใน

ทางที่ดี เช่นทำให้เกิดความร่วมมือและเพิ่มความสัมพันธ์ระหว่างประเทศเอเชียในด้านวิทยาศาสตร์และเทคโนโลยี นอกจากนี้ยังทำให้ผู้รับทุนชาวต่างประเทศรู้สึกเป็นหนี้บุญคุณประเทศไทย ที่ทำให้เขามีโอกาสได้ศึกษาในระดับปริญญาโท นอกจากนี้การที่สถาบันตั้งอยู่ในประเทศไทย ก็มีผลดีต่อประเทศไทยอย่างยิ่ง เพราะบุคลากรของสถาบันเป็นผู้มีความรู้ความสามารถ งานด้านวิจัยต่างๆ ของสถาบันสามารถช่วยในการพัฒนาประเทศไทยให้ก้าวหน้ายิ่งขึ้น สำหรับผู้รับทุนชาวไทยเกิดความภาคภูมิใจ รักพระเจ้าอยู่หัว และเกิดแรงผลักดันให้เกิดจิตสำนึกที่ดีต่อสังคม/สิ่งแวดล้อม และคิดจะพัฒนาบ้านเมือง ดังตัวอย่างคำพูดของผู้รับทุนดังนี้

ผู้รับทุนชาวไทย

“เป็นประโยชน์ต่อการพัฒนาประเทศ ทำให้นักศึกษาต่างชาติดีทัศนคติที่ดีต่อไทย”

“เป็นการสร้างสัมพันธ์ภาพกับเพื่อนบ้านที่ได้รับทุนได้รำลึกถึงพระมหากษัตริย์คุณ”

“เป็นการสร้างจิตสำนึกที่ดีต่อสังคมสิ่งแวดล้อมในประเทศไทยให้กับผู้รับทุนทางอ้อม เนื่องจากวิชาอาจารย์จะเน้นมาก เรื่องการพัฒนาแบบยั่งยืนตาม concept ของในหลวง”

“ทำให้รู้แหล่งวิชาการมากขึ้น ทำให้เกิดแนวคิดในการจะช่วย transfer technology จากต่างประเทศ”

ข้อเสนอแนะ

จากการติดตามประเมินแผนการอุดหนุนสถาบันระยะที่ 5 ระหว่างปี พ.ศ. 2533-2537 ปรากฏผลเป็นที่น่าพอใจในระดับหนึ่ง อย่างไรก็ตามจากผลการประเมินทำให้มีโอกาสเห็นช่องทางจะปรับปรุงทั้งระดับแผนและการดำเนินงานความร่วมมือระหว่างรัฐบาลไทยและสถาบันเทคโนโลยีแห่งเอเชียในระยะต่อไป เพื่อประกันคุณภาพและประสิทธิภาพของความสำเร็จ จึงน่าจะมีการพิจารณาที่เสนอดังต่อไปนี้

ผู้รับทุนชาวต่างประเทศ

“It surely improved the image of Thailand. Futhermore, the students who get the scholarship always feel grateful and indebted toward the Royal Thai Government”

“Not only materialistic but also spiritually, the schoolarship will encourage students to fullfill their task at AIT”

“Most students in AIT conduct their thesis and research in Thailand. So in the end, it is Thailand who will get the benefit first”

1. กระบวนการจัดทำแผนและงบประมาณจากการศึกษาการจัดทำแผนพบว่ามีลักษณะเป็นแผนการเงินที่ขอรับการสนับสนุนจากรัฐบาลไทยในช่วงระยะเวลา 5 ปี มากกว่าเป็นแผนพัฒนาและจัดทำตามนโยบายของสถาบันเป็นส่วนใหญ่ ทางรัฐบาลไทยจะเป็นผู้รับรองแผนที่สถาบันเสนอมาเท่านั้น บทบาทที่ควรจะเป็นคือ รัฐบาลไทยและสถาบันควรจะวางนโยบายร่วมกันจัดทำแผนความร่วมมือ (Cooperative plan) กล่าวคือ รัฐบาลไทยเป็นผู้กำหนดนโยบาย (Top-down) โดยพิจารณาความต้องการของรัฐบาลไทยและความต้องการของสถาบัน และให้มีการตกลงร่วมกันระหว่างรัฐบาลไทยและผู้บริหารของสถาบัน ก่อนที่สถาบันจะจัดทำแผน ในลักษณะนี้จะเป็นแผนที่สนองวัตถุประสงค์ของทั้งสถาบันและรัฐบาลไทย

2. ในด้านการจัดสรรงบประมาณ รัฐบาลไทยได้จัดสรรงบประมาณในหมวดเงินอุดหนุนและหมวดรายจ่ายอื่นให้แก่สถาบัน การจัดสรรหมวดเงินอุดหนุนที่เป็นการอุดหนุนทั่วไป (Block-grant) เป็นที่พอใจของสถาบัน ซึ่งการจัดสรรในลักษณะนี้ผู้รับการสนับสนุนสามารถบริหารงบประมาณของตนเองได้ และสามารถนำไปหาประโยชน์ทางอ้อม ซึ่งจะนำดอกผลมาเป็นรายได้ของสถาบัน ทำให้เกิดความคล่องตัวในการใช้จ่าย และเป็นการจัดสรรที่ดีกว่าการจัดสรรให้แก่หน่วยงานของรัฐ เช่น เงินอุดหนุนการดำเนินงาน เงินอุดหนุนสมทบค่าก่อสร้าง เงินทุนเล่าเรียนหลวง เป็นต้น แต่การจัดสรรในลักษณะนี้ จำเป็นต้องมีรายละเอียดของโครงการที่เสนอขออย่างชัดเจน ที่จะทำให้ผู้จัดสรรสามารถวิเคราะห์งบประมาณที่จะสนับสนุนให้อย่างเพียงพอ สามารถปฏิบัติงานได้และจำเป็นต้องมีระบบการรายงาน การติดตามประเมินผล ในกระบวนการบริหารงบประมาณที่รัฐบาลไทยให้การสนับสนุนด้วยจึงสมควรเพิ่มขึ้นตอนเหล่านี้ไว้ในกระบวนการงบประมาณด้วย

3. การจัดสรรงบประมาณสมทบค่าก่อสร้าง ยังมีปัญหาของการจัดสรร เพราะผลของการติดตามผลชี้ชัดว่า สถาบันไม่สามารถดำเนินการตามที่เสนอของงบประมาณได้ เนื่องจากงบประมาณที่ให้เป็นงบสมทบ สถาบันจำเป็นต้องหาผู้สมทบก่อสร้างจากประเทศอื่นด้วยเพื่อให้เพียงพอต่อการสร้างอาคารและประเทศเหล่านั้นต้องมีความเห็นสอดคล้องในการที่จะสร้างอาคารเดียวกัน การจัดสรรงบประมาณในลักษณะเล็กๆ น้อยๆ เช่นนี้ไม่ค่อยเห็นผล ดังนั้น การจัดสรรงบประมาณสมทบค่าก่อสร้างจึงอาจไม่เกิดประโยชน์เท่าที่ควร สมควรให้มีการจัดทำแผนระยะยาวของการสนับสนุนการก่อสร้างอาคาร โดยอาจเป็นแผนที่รัฐบาลไทยสนับสนุนการก่อสร้างอาคารหนึ่งหลังหรืออาจหางบประมาณสมทบจากภาคเอกชน จัดสรรเป็นการก่อสร้างอาคารในช่วง 5 ปี หรือ 10 ปี

4. เพื่อประสิทธิผลของแผนพัฒนา จึงเห็นควรที่สำนักงานคณะกรรมการการศึกษาแห่งชาติควรเพิ่มบทบาทในการเป็นผู้ประสานงาน การกำหนดนโยบายที่ตกลงร่วมกันระหว่างรัฐบาลไทยและสถาบัน เพื่อการจัดทำแผนพัฒนาระยะ 5 ปี เพิ่มบทบาทในการเป็นผู้กำกับติดตามการ

ประเมินผลและการรายงานของสถาบัน เพื่อให้เกิดการดำเนินงานและการใช้จ่ายงบประมาณที่มีประสิทธิภาพ ประสิทธิผล สอดคล้องกับนโยบายของรัฐบาลไทยและสถาบันให้มากที่สุด

5. จากการประเมินผลการเบิกจ่ายของสถาบันพบว่า มีการขอโอนหมวดรายจ่ายอื่นไปหมวดเงินอุดหนุนเป็นจำนวนมาก และมีงบประมาณจำนวนหนึ่งที่มีได้เบิกจ่าย อีกทั้งมีบางรายการที่สามารถเร่งรัดการเบิกจ่ายได้ เช่น ภาษีรายได้ประจำปี ควรจะทยอยการขอเบิกจ่ายมากกว่าที่จะขอเบิกจ่ายเป็นยอดเดียวตอนปลายปีงบประมาณ

ดังนั้น หากสถาบันเน้นบทบาทจัดทำแผนพัฒนาและแผนการเงินให้สอดคล้องกับกิจกรรมที่จะดำเนินการ และบริหารการเงินให้เป็นไปตามงบประมาณที่ได้รับจัดสรร ก็จะเป็นประโยชน์ทั้งฝ่ายรัฐบาลไทยและสถาบัน อีกทั้งสถาบันน่าจะจัดทำรายงานการดำเนินงานการใช้จ่ายเงินอุดหนุน (Accountability Report) เสนอสำนักงานคณะกรรมการการศึกษาแห่งชาติ

6. จากผลการศึกษาค่าใช้จ่ายต่อนักศึกษา สถาบันมีค่าใช้จ่ายต่อนักศึกษาสูง (ประมาณ 475,000 บาท/นักศึกษา/ปี) เมื่อเทียบกับค่าใช้จ่ายในการผลิตบัณฑิตภายในประเทศ แต่ถ้าเทียบกับกรณีที่รัฐบาลไทยต้องส่งนักศึกษาไปศึกษาต่างประเทศแล้ว จะเห็นว่าถูกกว่า ดังนั้น หากรัฐบาลมีนโยบายเพิ่มจำนวนทุนและสาขาที่ประเทศต้องการ หรือให้สถาบันเพิ่มรับคนไทยเข้าศึกษาแล้วจะเป็นประโยชน์ต่อการศึกษาในประเทศเป็นอย่างมาก

7. จากผลของการศึกษาค่าใช้จ่ายต่อนักศึกษาของสถาบันพบว่า ค่าใช้จ่ายต่อนักศึกษาที่สถาบันจะต้องใช้นั้น มีจำนวนสูงกว่าทุนการศึกษาที่สถาบันได้รับจากหน่วยงาน/องค์กรต่างๆ เพื่อให้สอดคล้องกับงบประมาณที่สถาบันต้องรับผิดชอบในการผลิตนักศึกษา สถาบันอาจใช้ข้อมูลเหล่านี้ในการปรับค่าใช้จ่ายต่อทุนการศึกษาที่เรียกเก็บจากหน่วยงานต่างๆ โดยเฉพาะหน่วยงานที่สนับสนุนเฉพาะทุนการศึกษาเท่านั้น

8. จากการศึกษพบว่า มีปัญหาของข้อมูลด้านการเงินที่เกิดจากปีงบประมาณของสถาบัน (กรกฎาคม-มิถุนายน) และปีงบประมาณแผ่นดิน (ตุลาคม-กันยายน) ไม่ตรงกัน ซึ่งทางสถาบันจะมีการเก็บข้อมูลและการรายงานตามปีงบประมาณของสถาบันเท่านั้น ดังนั้นในการศึกษาจึงเป็นการยากแก่การจัดเก็บข้อมูลและการเปรียบเทียบ ทางสถาบันควรมีการรายงานทั้งสองรูปแบบคือ ถ้าเป็นงบประมาณของรัฐบาลไทยควรมีรายงานการเบิกจ่ายตามปีงบประมาณแผ่นดินด้วย

9. เพื่อให้สอดคล้องกับวัตถุประสงค์ของการให้ทุนเล่าเรียนหลวงว่าต้องการให้แก่ นักศึกษาประเทศไทยและประเทศที่กำลังพัฒนาในภูมิภาคเอเชีย ดังนั้น รัฐบาลไทยควรจะได้กำหนดเงื่อนไขของการให้ทุนว่า ต้องการให้แก่คนไทยร้อยละเท่าไร ชาวเอเชียในประเทศที่กำลังพัฒนาอื่น ๆ

ร้อยละเท่าไรให้ชัดเจน โดยแบ่งสัดส่วนของทุนให้เท่าๆ กันในแต่ละภาคการศึกษา เพื่อจะได้มี นักศึกษาทุนเล่าเรียนในสาขาต่างๆ ใกล้เคียงกัน และควรกำหนดด้วยว่าต้องการให้ทุนแก่ผู้ที่มีผล การเรียนดีเยี่ยมเช่นเป็นผู้ที่ได้เกียรตินิยมในระดับปริญญาตรีหรือเป็นผู้ที่มีผลการเรียนดีที่สุด 10 อันดับแรกของผู้ที่ขอทุนภายในของสถาบัน ทั้งหมด หรือเป็นผู้ที่มีผลการเรียนในอันดับที่ 1 ของชั้น หรือของสถาบันการศึกษาที่จบปริญญาตรี เพื่อจะได้สอดคล้องกับที่ได้กำหนดไว้ว่า ผู้ที่ได้รับทุนเล่า เรียนหลวงหมายถึงผู้ที่มีผลการเรียนดีเยี่ยม และเป็นรางวัลเกียรติยศซึ่งนำมาซึ่งความภาคภูมิใจแก่ผู้ ที่ได้รับ โดยไม่ต้องกำหนดสาขาเพื่อว่าทุนเล่าเรียนหลวงจะได้คนที่มีความสามารถดีเยี่ยม และได้ เรียนในสาขาที่อยากเรียนและจะได้นำประโยชน์ไปพัฒนาประเทศ

10. ในใบสมัครของสถาบัน ควรมีการระบุชื่อทุนเล่าเรียนหลวงให้นักศึกษาเลือกขอทุนมา โดยเฉพาะ และควรมีสิ่งจูงใจในการขอทุน เช่น มีใบประชาสัมพันธ์ให้ทราบว่าทุนเล่าเรียนหลวงให้ ทุนในส่วนค่าใช้จ่ายส่วนใดบ้างซึ่งอาจจะเพิ่มทุนในด้านให้ค่าใช้จ่ายในการทำวิทยานิพนธ์หรือทำวิจัย และมีการเข้าเฝ้ารับพระราชทานทุนจากพระบาทสมเด็จพระเจ้าอยู่หัวทุกปี เพื่อจะได้จูงใจให้ผู้ที่มี ผลการเรียนดีเยี่ยมมาสมัครขอทุนเล่าเรียนหลวงมากๆ จะได้มีโอกาสในการคัดเลือกผู้ที่มีผลการ เรียนดีเยี่ยมสูงขึ้น อีกทั้งควรจะมีการประชาสัมพันธ์เกี่ยวกับทุนเล่าเรียนหลวงโดยเฉพาะ เจตนารมณ์ของทุนที่มุ่งพัฒนาประเทศและให้ชนชาวมุสลิมภาคเอเชียอยู่กันอย่างสันติ

11. ควรจะมีการรายงานผลการเรียนของนักศึกษาทุนเล่าเรียนหลวงให้แก่ทางสำนักงาน คณะกรรมการการศึกษาแห่งชาติได้ทราบทุกภาคการศึกษา และอาจจะมีรางวัลให้สำหรับนักศึกษา ที่มีผลการเรียนดีเยี่ยม เพื่อจูงใจในการสมัครขอรับทุนเล่าเรียนหลวง

12. ควรมีการจัดสรรค้ำระหว่างผู้ประสานงานทุนเล่าเรียนหลวงทุกหน่วยงาน และนักศึกษา ที่ได้รับทุนเล่าเรียนหลวงทุกคน เพื่อสร้างความสัมพันธ์อันดีระหว่างนักศึกษาและทำให้นักศึกษา ชาบซึ่งและประทับใจในประเทศไทย

13. ทางสถาบัน ควรมีเกณฑ์มาตรฐานที่ชัดเจนในการเทียบผลการเรียนของนักศึกษา จากแต่ละประเทศให้เป็นมาตรฐานเดียวกัน เพื่อให้การคัดเลือกนักศึกษาเข้ารับทุนต่างๆ เป็นไปใน มาตรฐานเดียวกัน

14. สำหรับทุนเล่าเรียนหลวงอีก 8 ทุนที่ทางสถาบันค้างอยู่ ควรจะได้จัดสรรให้ในปีต่อๆ ไปรวมทั้งทุนที่ให้แล้วเรียนไม่จบอีก 1 ทุน อีกทั้งในวโรกาสแห่งปีกาญจนาภิเษกสมโภช เห็นสมควร ให้เพิ่มจำนวนทุนเล่าเรียนหลวง เช่น เพิ่มขึ้นอีกอย่างน้อยปีละ 5 ทุน ตลอดจนน่าจะพิจารณาให้ ทุนแก่ผู้สำเร็จปริญญาตรี หรือโท ที่มีผลการเรียนดีเยี่ยมได้ศึกษาต่อในระดับปริญญาเอกด้วย

15. ผู้รับทุนเล่าเรียนหลวงบางส่วนจำเป็นต้องใช้เงินในการทำวิทยานิพนธ์ การให้ทุนตามระบบเดิมจึงควรมีการยืดหยุ่นเพื่อนำเงินมาใช้ในการทำวิทยานิพนธ์ หรือเพิ่มเงินทุนสำหรับการทำวิทยานิพนธ์โดยเฉพาะ และควรมีการพิจารณาให้รางวัลวิทยานิพนธ์ดีเด่น ในบรรดาผู้รับทุนเล่าเรียนหลวงเพื่อเป็นการส่งเสริมให้ผู้รับทุนทำวิทยานิพนธ์ที่มีคุณภาพ นอกจากนี้ผู้ได้รับรางวัลวิทยานิพนธ์ดีเด่นควรได้รับการส่งเสริมให้นำเสนอผลงานในต่างประเทศ หรือพิมพ์ในวารสารต่างประเทศ เพื่อเป็นการเผยแพร่ชื่อเสียงของทุนเล่าเรียนหลวง

16. จากการศึกษาผู้รับทุนเล่าเรียนหลวงนั้น ส่วนใหญ่ทำงานในภาคเอกชนซึ่งทำให้กำลังสมองไปสนับสนุนในภาคเอกชนอย่างดี จึงน่าจะมีการพิจารณาว่าจะมีแนวทางใดที่จะเป็นการทำให้ทุนเล่าเรียนหลวงนี้ส่งเสริมคนในภาครัฐบาลให้มากขึ้น เพื่อเสริมสร้างกำลังคนที่มีความรู้ ความสามารถในภาครัฐให้มากขึ้น

17. สถาบันน่าจะจัดระบบฐานข้อมูลเกี่ยวกับการใช้จ่ายเงินอุดหนุน และเกี่ยวกับกระบวนการบริหารเงินทุน ตลอดจนข้อมูลเกี่ยวกับผู้รับทุนเล่าเรียนหลวง อีกทั้งจัดทำรายงานเชิงประเมินตนเอง (self-assessment report) เสนอสำนักงานคณะกรรมการการศึกษาแห่งชาติเป็นประจำทุกปี

18. สำนักงานคณะกรรมการการศึกษาแห่งชาติ น่าจะจัดให้มีการติดตามประเมินความร่วมมือระหว่างรัฐบาลไทยและสถาบันเทคโนโลยีแห่งเอเชียในส่วนของรัฐบาลไทยให้การสนับสนุนเป็นประจำทุกแผน ดังเช่นในการติดตามประเมินแผนการอุดหนุนในระยะที่ 5 โดยกำหนดให้มีการประเมินถึงกลางแผน และการประเมินเมื่อสิ้นสุดแต่ละแผนด้วย สำหรับรูปแบบการติดตามและประเมินแผนนั้นสามารถจะใช้รูปแบบการติดตามและประเมินแผนการอุดหนุน 5 ปีระยะที่ 5

เอกสารอ้างอิง

คณะกรรมการการศึกษาแห่งชาติ, สำนักงาน; “ผลการประชุมคณะอนุกรรมการพิจารณาให้การอุดหนุนด้านงบประมาณของสถาบันเทคโนโลยีแห่งเอเชีย ปี 2531” และ “มติคณะกรรมการการศึกษาแห่งชาติ ในคราวประชุมครั้งที่ 4/2531 เมื่อวันที่ 15 เมษายน 2531” (เอกสารฉบับอัดสำเนา); พ.ศ 2531.

เทียนฉาย กิระนันท์ *การวิเคราะห์ค่าใช้จ่ายของบัณฑิตศึกษา จุฬาลงกรณ์มหาวิทยาลัย* กรุงเทพมหานคร บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 2534.

เลขาธิการคณะรัฐมนตรี, “สำนักงาน ; หนังสือที่ นร 0202/9290 ลงวันที่ 3 สิงหาคม 2531 เรื่อง การให้ความสนับสนุนด้านงบประมาณของสถาบันเทคโนโลยีแห่งเอเชีย สำหรับปี 2533-2537 แจ้งมติคณะรัฐมนตรี” (เอกสารฉบับอัดสำเนา) พ.ศ. 2531,

สถาบันเทคโนโลยีแห่งเอเชีย (AIT), “โครงการความช่วยเหลือร่วมมือระหว่างรัฐบาลไทยและสถาบันเทคโนโลยีแห่งเอเชีย ปีงบประมาณ 2538-2542 “เอกสารฉบับอัดสำเนา” ปี 2536. สมหวัง พิธิยานุวัฒน์ และคณะ. *การประเมินโครงการพัฒนาและส่งเสริมผู้มีความสามารถพิเศษทางวิทยาศาสตร์และเทคโนโลยี (พสวท.) ระยะที่ 2* มกราคม 2539.

สมหวัง พิธิยานุวัฒน์ และคณะ. *การศึกษาแนวโน้มความต้องการและการผลิตกำลังคนระดับสูงกว่าปริญญาตรี*, กันยายน 2539.

Asian Institute of Technology. *1990 AIT Annual report*. Thailand: The Office of Media and Information Services, 1991,

Asian Institute of Technology. *1991 AIT Annual report*. Thailand: The Office of Media and Information Services, 1992,

Asian Institute of Technology. *1992 AIT Annual report*. Thailand: The Office of Media and Information Services, 1993,

Asian Institute of Technology. *1993 AIT Annual report*. Thailand: The Office of Media and Information Services, 1994,

Asian Institute of Technology. *1994 AIT Annual report*. Thailand: The Office of Media and Information Services, 1995,

AIT; *Prospectus 1996*, Asian Institute of Technology, May 1995, Bangkok, Thailand.

AUTHORS

Susan M. Brookhart, Ph.D.	<i>Duquesne University</i>
Pornthip Chaiyaso, Ph.D.	<i>Kasetsart University</i>
Boonreang Kajornsin, Ph.D.	<i>Kasetsart University</i>
Samnao Kajornsin, Ph.D.	<i>Kasetsart University</i>
Nikom Nak-ai, M.Ed.	<i>The Office of National Primary Education Commission</i>
Somwung Pitiyanuwat, Ph.D.	<i>Chulalongkorn University</i>
Wilaiwan Pitiyanuwat, M.Ed.	<i>Chulalongkorn University</i>
Panurut Rattayapas	<i>The Office of the National Education Commission</i>
Auyporn Ruengtrakul, M.Ed.	<i>Chulalongkorn University</i>
Tasanee Sahawat	<i>The Asian Institute of Technology</i>
Sopana Takaew	<i>The Office of the National Education Commission</i>
James S. Terwilliger, Ph.D.	<i>University of Minnesota</i>
Wantani Thaitiang	<i>The Office of the National Education Commission</i>

พิมพ์ที่โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย โทร. 2153626
นายประเสริฐ ศิลป์พัฒน์ ผู้พิมพ์ผู้โฆษณา ตุลาคม 2540

4012-276/1,000(2)

