ABSTRACT
This study tries to investigate the importance of Daniel Stufflebeam’s Context-Input-Process-Product or CIPP Model in academic performance of foreign students based on a comprehensive review of the literature. Moreover, Program evaluation can include any or a variety of at least 35 different types of evaluation, such as for needs assessments, accreditation, cost/benefit analysis, effectiveness, efficiency, formative, summative, goal based, process, outcomes, and a lot more. In this particular study, the researcher undertakes the program theory that focuses on the program’s design and concept. The program theory breaks down the components of the program and shows anticipated short term and long term effects. A real analysis of the program theory examines how the program is organized and how that organization will lead to desired outcomes. It will also reveal unintended or unforeseen consequences of a program, both positive and negative. This review can make a unique view of program evaluation of the English Language Proficiency Programs models.

KEYWORDS: Stufflebeam’s CIPP Model & Program Theory, Program evaluation

STUFFLEBEAM’S CONTEXT-INPUT-PROCESS-PRODUCT MODEL
The most important contribution to a decision management oriented approach to educational evaluation has been presented by STUFFLEBEAM, DANIEL (2004). His approach is recognized as the context, input, process and product or CIPP model. This comprehensive model considers evaluation as a continuing process. Information is given to the management for the purpose of decision making. It is a three-step process: delineating the information necessary for collection, obtaining the information and providing the information to interested parties. Context evaluation involves studying the environment of the program and the purpose is to define the relevant information, focus on unmet needs and missed opportunities, and diagnose the reasons for unmet needs. It is actually a way to provide information and determine how to utilize resources to meet program goals. It evaluates specific aspects of the program. Process evaluation addresses the implementation decisions that control and manage the program. Product evaluation is the gathering of data to determine the extent the objectives are being attained. It provides evaluators with information that enable them to decide whether to continue, terminate or modify the program. The model can be illustrated in Figure 1 below.
Table 1: The CIPP (Context-Input-Process-Product) Model of Evaluation

<table>
<thead>
<tr>
<th>Aspects of Evaluation</th>
<th>Types of Decisions</th>
<th>Kind of Questions Answered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context Evaluation</td>
<td>Planning Decisions</td>
<td>What Should We Do?</td>
</tr>
<tr>
<td>Input Evaluation</td>
<td>Structuring Decisions</td>
<td>How Should We Do It?</td>
</tr>
<tr>
<td>Process Evaluation</td>
<td>Implementing Decisions</td>
<td>Are We Doing It as Planned?</td>
</tr>
<tr>
<td>Product Evaluation</td>
<td>Recycling Decisions</td>
<td>Did It Work?</td>
</tr>
</tbody>
</table>

The Program Impact Theory


- The program theory is composed of the impact theory and the process theory.
- The impact theory is the theory showing the results or outcomes that are supposed to be produced by the program activities. These outcomes include both proximal outcomes (more immediate or direct results) and distal outcomes (more ultimate or longer term outcomes that result from the proximal outcomes). The action theory (or action hypothesis) specifies the expected immediate or proximal outcomes of the program actions; the conceptual theory (or conceptual hypothesis) specifies the expected distal or longer term outcomes.
- The next two components of the program theory (service utilization and organizational plan) are known as the program process theory and they require process evaluation.
- The service utilization plan is the expectations about how program clients will become engaged with the program and participate in program activities. It describes the program-target transactions (and what is behind those transactions) from the perspective of the program target participants. It explains what the participants must do if they are to be part of the program.
- The organizational plan is the expectations about how the program will be organized and maintained. It describes the program-target transactions (and what is behind those transactions) from the perspective of program management. It explains what the organization must do to deliver and maintain an effective program.

There are three interrelated components of the Program Theory and all these components compliment each other. The inputs and activities represent the program organization. The outputs represent the service utilization. The process theory includes both the program organization and the service utilization. Next is the impact theory (labeled Outcomes) which includes the initial outcomes, intermediate outcomes, and the longer term outcomes. These can be illustrated in Figure 1 shown below.
Figure 1: The program theory and its three components

It should be obvious that there are a lot of things that need to be examined when one is articulating or describing a program theory (CHEN, 2002). In articulating (or describing) a program theory, first, there is a need to come up with a clear definition of the program, its objectives, and its boundaries that are negotiated and agreed upon by the evaluator and the primary stakeholders.

Next, one can use some or all of the following procedures for articulating or explicating or describing the program theory (ROSSI, 2004). Articulation is an iterative and successive process, where one obtains information, listens to stakeholders’ views and reactions to the developing program theory, obtains more information and listens to stakeholders’ reactions to the revised program theory, and continues this process until there is some agreement on the description of the program theory (i.e., how the program is intended to operate). The following approaches are suggested:

- Review program documents and other secondary or extant data. A few examples are written descriptions of the program, the authorizing legislation for the program, grants, annual reports, brochures explaining the program, mission statements, lists of program goals and objectives, program manuals, and job descriptions.
- Interview program personnel and other stakeholders. One should interview clients of the program (i.e., people who have completed the program services), other members of the target population who have not yet completed the program services, program staff and management, funding agencies and program sponsors, and anyone else with information about the conceptualization and operation of the program. It is important to obtain and listen to the different
perspectives of these information providers and try to come up with an agreed upon description of the program as it is intended to operate.

- Make site visits and observe the program. The strength of collecting data through observation is that one can see what people are doing, rather than having to only rely on what they say they do. During observation, one can “explore the program reality firsthand.” This process also allows one (the evaluator) to provide an independent perspective on what is done in the program. Look for any discrepancy between the
  - program theory being described up to this point (what is supposed to be done) and the program reality being observed (what is actually done).

Furthermore, it is suggested that during the data collection expedition, one should explore the following topics:

- Determine the program goals and objectives. Rather than directly asking program personnel and stakeholders “What are the objectives of your program?” ask about the expected outcomes and then work your way back to the objectives. Identify impact theory goals and objectives (e.g., program participants will pass a test on drug knowledge; program participants will improve their communication skills; program participants will not resume drug use after leaving the program). Also, identify process theory goals and objectives. Throughout the process of data collection, one should be on the lookout for possible side effects or unintended outcomes that may result from the program (ROSSI, 2004).

- Determine the program functions, components, and activities. The focus here is primarily on describing the intended program process. In other words, what is going to be done to bring about the expected outcomes? Examine each of the components of the service utilization and organizational plans, and list the activities participants must engage in (ROSSI, 2004).

- Determine the logic or sequence that links the program functions, activities, and components. How are the various program components and activities temporally sequenced? What program activities require special coordination? What is the logic of the program impact theory and what kinds of conceptual sequencing does it imply. What is the logic model of the program? Are the activities and outcomes measurable? What successful criteria or standards will apply to the outcomes (ROSSI, 2004).

**REVIEW OF RELATED LITERATURE**

Rutman (1984) defines a program as an intervention or set of activities mounted to achieve external objectives, that is, to meet some recognized social needs, or to solve an identifiable problem. Typically, organizations work from their mission to identify several goals which must be reached to accomplish their mission. What is program evaluation? It is a process of carefully collecting information about a program or some aspects of a program in order to make necessary decisions about the program.

Program evaluation can include any or a variety of at least 35 different types of evaluation, such as for needs assessments, accreditation, cost/benefit analysis, effectiveness, efficiency, formative, summative, goal based, process, outcomes, and a lot more. The type of evaluation one undertakes to improve his program depends on what he wants to learn about the program. Program evaluators may assess programs on several dimensions to determine whether the program works. Rossi et al.
(2004) divide these dimensions into five categories: needs assessment, program theory, process analysis, impact analysis, and cost-benefit & cost-effectiveness analysis. When an evaluator examines the nature of the problem that the program is meant to address, it is called needs assessment. This would include evaluating who is being affected by the problem, how widespread the problem is, and what effects stem from the problem. The program theory is the formal description of a program’s concept and design. The program theory breaks down the components of the program and shows anticipated short term and long term effects. A real analysis of the program theory examines how the program is organized and how that organization will lead to desired outcomes. It will also reveal unintended or unforeseen consequences of a program, both positive and negative.

The third category is process theory which looks beyond the theory of what the program is supposed to do, and instead evaluates how the program is being implemented. The evaluation determines whether the target population is being reached, people are receiving the intended services, staff are adequately qualified, and many others. The fourth category is impact evaluation which determines the causal effects of the program.

Finally, cost-benefit and cost-effectiveness analysis assesses the efficiency of the program. Evaluators outline the benefits and the cost of the program for comparison. The most difficult part of evaluation is determining whether the program itself is causing observed impacts. Events or activities outside the program may be the real cause of the observed outcome. Causation is difficult to determine. People select themselves to participate in the program. Those who do participate may differ from those who do not in important ways. They may become determined to find a job or have better support resources.

A study titled “The impact of English-language proficiency on international graduate students' perceived academic difficulty”, the author, Ming Xu, presents an empirical test of the impact of English proficiency as indicated by student self-ratings and TOEFL scores, and selected nonlinguistic variables on the level of academic difficulties perceived by international graduate students. A simple random sample of 450 international graduate students from three large universities in upstate New York were contacted by mail in the fall of 2008. 245 completed the Survey of International Students' Academic Language Needs. It was found that English Proficiency in ESL programs and school curriculum designs were the most significant predictors of the perceived level of academic difficulty.

In another study conducted by A.A. Winke, and published as a journal article in “Higher Education, 2009” Volume 31, no 3, 325-340, it investigated associations between studying in a foreign English language and students' academic achievement and behaviour. Data were gathered in the Netherlands, at the Delft University of Technology (DUT). The results show that the success rate of foreign students was about the same as that of the Dutch students, but that foreign students needed more time to pass the examination. Language proficiency played a major part in explaining differences in study achievement as did the presence or absence of a selection procedure in the student's country of origin. Furthermore, the results show that the study behaviour of foreign students was significantly different from the Dutch students: their average grades on examinations
were lower, they made more attempts before passing an examination, they were more likely to postpone examinations and to follow a different order from the recommended one. Suggested measures to improve foreign students' study success include requiring higher levels of language proficiency and stricter selection of foreign students in the country of origin. It is suggested that a system of mentors which provides for intensive coaching of foreign students in their first year may prevent a slow start.

In the Netherlands, internationalization of higher education led to an increasing use of a foreign language—for example English—as the medium of instruction. This raised the issue whether learning in a foreign language affects academic performance, given the language problems international students are faced with when studying abroad. This article reviews research in this field and presents a study into the academic achievement of 90 Indonesian engineers. These subjects took a one-year English-medium graduate course at the International Institute for Hydraulics and Environmental Engineering in Delft. Language proficiency was measured by TOEFL and academic success was defined as the average score on seven written examinations. A most remarkable finding is the observation of a cut-off point in the relationship between English proficiency and academic success. This has led to the hypothesis that there is a range of TOEFL scores within which a better command of English increases the chance of being academically successful. The lower and upper limits of this range may vary according to academic discipline or may even be institution-specific. This study was published as a journal article: “Career Placement Concerns of International Graduate Students: A Qualitative Study” by Yih-Jiun Shen for Higher Education Vol. 32, 2008.

In the Philippine context, all member-schools and institutions of Philippine Association of Graduate Education or PAGE strictly rule admission requirements for international students, and set high standards of academic performance of graduate students as part of its thrust for continuing professional education and upgrading. PAGE’S main theme “Dissemination and Utilization of Research Outputs,” for its 2011 Annual Assembly was articulated by its keynote speakers, CHED Commissioner Nona Ricafort and Dr. Isagani Cruz who lectured about an impressive academic performance of international students pursuing graduate programs in PAGE schools and institutions over the years.

**DISCUSSION**

Daniel Stufflebeam (2004), who advocated his CIPP model of program evaluation encourages all evaluators that this millennium is an opportune time to critically appraise their program evaluation approaches and decide which ones are most worthy of continued application and further development. It is also important that evaluators should identify approaches that need to be abandoned. In this context, he identifies and assesses twenty two (22) approaches often employed to evaluate programs. These approaches in varying degrees are unique and comprise most evaluation efforts. Stufflebeam uses the term *evaluation approach* instead of *evaluation model*, because the former is broad enough to cover illicit as well as laudatory practices. He gives a broader view of program and evaluation. It encompasses evaluations of any coordinated set of activities directed at achieving goals.
CONCLUSION
The study of these evaluation approaches is vital not only to professionalize program evaluation but also to produce its scientific advancement and operation. Scientifically, such a review can help evaluation researchers identify, examine, and address conceptual and technical issues pertaining to the discipline. Operationally, a critical view of alternatives can help evaluators consider and assess optional frameworks for planning and conducting particular studies. On this point, Stufflebeam has found out that different approaches may work differentially well, depending on the evaluation’s context. Often, it is advantageous to borrow strengths of different approaches to create a “best fit” approach for specific evaluation projects. Therefore, it is advisable that evaluators should develop a repertoire of different legitimate approaches they can use, plus the ability to discern which approaches work best under what circumstances. However, the main value in studying alternative approaches is to discover their strengths and weaknesses, decide which ones merit substantial use, determine when and how they are best applied, and obtain direction for improving these approaches and devising better alternatives.

REFERENCES