Evaluating Training Effectiveness: Focus on the Training of Public Enterprise Managers

GABRIEL U. IGLESIAS*

Evaluation of training effectiveness is generally accepted as important in measuring levels or changes in individual, unit and organizational performance. A comparison of the experience of selected developed and developing countries, however, show that there has been relative neglect, on the part of the latter, to use more rigorous and reliable evaluation instruments, techniques, and methodologies in measuring the effectiveness of training conducted on public enterprise managers. There is need therefore to strengthen the capability of training institutions in evaluating training effectiveness.

Introduction

Although an enormous amount of money and time is spent by organizations in training personnel at various levels, evaluation of training is a very neglected aspect of training both in the developed and developing countries. An interregional seminar on public enterprises training reported that "generally, the existing evaluation of training tends to emphasize transfer of knowledge during the training process to the general neglect of evaluating the effect of training on the performance of the trainee in the work situation."[1] Although the focus of this paper is on evaluation of training

of public enterprise managers, it has also general relevance to evaluating training effectiveness of managers in the public service.

There appears to be a general agreement that evaluation of training effectiveness is an important tool in determining whether the use of organizational resources in training yield benefits in terms of improving performance at the individual, unit, and organization levels and in reducing or solving organizational and management problems. However, because it is generally more difficult to conduct a rigorous, more scientific type of evaluation mentioned above, there is a tendency to limit evaluation of training effectiveness to eliciting subjective evaluation on the value of training from participants, what is generally known as "reaction evaluation." For example, a survey made in the United States and Canada revealed that 85 (78%) out of 110 firms which evaluated their training used reaction evaluation, compared to 56 which tried to determine learning that took place, 59 which attempted to measure changes

* Professor and Director, Policy Studies Program, College of Public Administration, University of the Philippines.

[1] Syndicate A, "Report on Efficiency and Training in Public Enterprises," Interregional Seminar on Training Management in Public Enterprise held in Ljubljana, Yugoslavia, 29 September-10 October 1980, under the auspices of the International Center for Public Enterprise in Developing Countries (ICPE).

in the on-the-job behavior, and 50 which tried to evaluate results (i.e., changes in unit or organizational performance).[2] Similarly, in a developing country like the Philippines, evaluation of training generally emphasized reaction of participants to the training, particularly in eliciting their opinions on "substance training resources and methods."[3]

In developing countries, there has been relative neglect in evaluating training effectiveness which measures levels or changes in learning, in terms of knowledge, skills, and attitudes (immediate outcomes), in terms of changes in on-the-job behavior and performance (intermediate outcomes), and in terms of changes at the unit or organizational levels (ultimate outcomes).[4] This neglect may be attributed to the following: (1) lack of technical expertise within the training organization; (2) the greater investment in cost and time; (3) absence of appreciation and/or demand for evaluation from within the organization or the clientele; and (4) the "evaluation paradox" where "trainers trying to measure the success of the program are at the same time saying 'what the program plans to accomplish can't be measured.' "[5]

The difficulty of conducting the more technically rigorous types of evaluation, according to a regional workshop which evaluated training packages for public enterprise managers, "lies in that there must be the ability to measure the change [improvement] in the performance in the work situation of the trainee-manager and the ability to directly attribute this change in performance to the training undergone by the manager."[6]

The Need to Evaluate Training Effectiveness

Various interregional and regional workshops convened by ICPE have drawn attention to the urgent need to develop the capability of training institutions and public enterprises in developing countries for evaluating training effectiveness, particularly in evolving suitable evaluation, instruments, methodologies, and techniques. One workshop recommended the development of "more precise methods and techniques of evaluating public enterprise training."[7] In another meeting organized by ICPE in

[2] Ralph Centanello and Donald Kirkpatrick, "Evaluating Training Programs, in D. Kirkpatrick (Compiler), *Evaluating Training Programs* (Madison, Wisconsin: American Society for Training and Development 1975), pp.258-259.

[3] Philippine Executive Academy, *An Evaluation of Its Program* (Manila: Philippine Executive Academy, 1969), p. ii.

[4] The terms "immediate," "intermediate," and "ultimate outcomes" are types of outcome evaluation formulated by Peter Warr, Michael Bird, and Neil Rackham, *Evaluation of Management Training*, Second Edition (London: Gower Press, 1971).

[5] Dorothy Fast, "A New Approach to Quantifying Training Program Effectiveness," in Kirkpatrick, *op.cit.*, p. 148.

[6] "Final Intergrated Report," Regional Workshop on the Evaluation of Training Packages for Public Enterprise Managers, Bangkok, Thailand, 10-14 August 1981, under the auspices of the ICPE and UN Asian and Pacific Development Centre (APDC).

[7] "Draft Final Report," Interregional Seminar on Training Management Ljubljana, Yugoslavia, 29 September-10 October 1980 under the sponsorship of ICPE, p. 13.

the Caribbean, the group "emphasized that most conventional training programmes contain few, if any, built-in techniques to measure their effectiveness, and is usually quite vague about the 'depth' and intent of the learning."[8]

In summing up research findings and conclusions from various regional and interregional workshops and seminars on training and management of public enterprises, Mozina noted that:

> In the developing countries various attempts have been made so far to improve the valuation process; however, current systems of evaluation are still inadequate. The development of suitable methodologies for evaluating the contribution of training is a field which needs further study.[9]

Purpose and Objectives

The main purpose of this paper is to report on the experiences of selected developed and developing countries on the subject of evaluating training effectiveness which would, hopefully, be useful to those concerned with this activity — particularly trainers, academic researchers, civil service and public enterprise managers. An attempt was made to incorporate in this paper selected approaches, methodologies, and instruments used in the evaluation of training, some of which would be more suitable to the stage of development, capability, and resources of civil

service and public enterprise training institutions in developing countries.

Efforts will be especially directed at evolving practical and less costly evaluation instruments for some areas of evaluation which may not be as scientifically rigorous as those employed in advanced countries but which offers a useful and, to a certain extent, systematic approach to this problem from the viewpoint and the experience of developing countries. A compendium of the experiences of both developed, mainly American, and developing countries, mainly Philippine, in evaluating training effectiveness would provide the reader not only a convenient source book of the evaluation methods and instruments used and the types of evaluation being done, but also useful suggestions in guiding evaluators in selecting feasible, practical, and less costly alternative evaluation strategies and instruments.

Framework for Evaluating Training Effectiveness

"Evaluation of training effectiveness" connotes many meanings and different perceptions of usefulness to different people. To trainors and staff of training institutions, it could mean "finding out whether their program is serving the needs of their clientele and providing a useful though subjective feedback mechanism to guide them in improving their program in terms of course content, learning materials, methods of instruction, subjects to stress, facilities, duration, and others."

To public enterprises, the civil service government planners, budgeting, auditing, and officials of training institutions, evaluation of training serves many purposes, especially if it yields findings on how training led to improvement in individual behavior

---

[8] "Final Report," Regional Workshop on Management Training in the English-Speaking Caribbean, Bridgetown, Barbados, 13-18 October 1980 under the auspices of ICPE and the Caribbean Centre for Development Administration, p. 28.

[9] Speech of Stave Mozina, Regional Workshop on Management Training and Development in Public Enterprises, Karachi, Pakistan 10-15 January 1981, under the auspices of the ICPE and the Pakistan Institute of Management.

and performance on the job — for example, increased ability to plan, monitor, and appraise projects, increased capacity for interpersonal relations and communication. Increased capability for evaluation could also provide a basis for justifying the cost of training in financial terms and man-hours lost, to improve the process of allocating financial, manpower and physical resources to training, for designing more effective training programs, and as a tool to improve planning and management of the public service, particularly public enterprises.

It should be stressed, however, that "results" evaluation requires rigorous and complicated techniques which may not be within the capability — in terms of technical expertise and resources — of most public enterprises in developing countries.[10]

As noted earlier, even in advanced industrialized countries, evaluation of results or of ultimate outcomes which attempts to relate training directly to changes in unit or organizational performance is not often done because it is methodologically difficult. This is particularly so in evaluating managerial training, because other variables in the work situation tend to affect unit or organizational performance since the organization is open to many external influences and it may be difficult to ascribe to training alone changes in performance at the organizational level. It is doubtful whether

---

[10] Donald L. Kirpatrick's evaluation framework used the following types of evaluation: reaction, learning (immediate outcome in Warr, Bird, and Rackham's formulation), behavior (intermediate outcome), and results (ultimate outcome). See "Techniques in Evaluating Training Programs," in Kirkpatrick, *op.cit.*, pp. 1-14.

many civil service agencies and public enterprises in developing countries would have the capability for this type of evaluation so that this paper will not give as much stress on results or ultimate outcome evaluation.

Although there is generally universal acceptance of the value of training, especially for skills training as well as management training for managers and supervisory personnel, the acceptance of its intrinsic worth does not obviate the need to determine with some degree of precision its contribution to the performance of the trainees, the unit, and the organization. For those engaged in providing training, it is important that they have a measure of how well they are doing their function not only to further improve their programmes but also to justify their *raison d'etre* to both management and organizations where their trainees originate.

"Evaluation of training effectiveness," in its formal sense, is "the process of determining whether the general or specific goals or objectives of an organizational activity (training) have been attained and to what degree have their goals or objectives been achieved." Corollary to this is the notion that achievement of training goals or objectives are based on factors and conditions which influence such achievement. Evaluation of training effectiveness is an attempt to prove a causal relationship between training goals or objectives, the activity carried on to achieve them (training), and the effects or results (performance) of that activity.

Conceptual Problems in Effectiveness Evaluation

As a basic social process by which individuals, groups, and organizations

make judgment on the worth or value of an object or activity, evaluation is and can be done without following rigid standards or procedures of presenting objective evidence as a basis for judgment since this is part of a general process of assessment or appraisal. Statements, such as "he is good," "that ministry is providing effective services," and "she is an efficient manager" could be asserted and, proven to some extent, but scientific proof may not be possible unless both the appraiser and the appraised could agree on the criteria or standards used to measure "good," "effective," and "efficient."

Thus, evaluation of training effectiveness as a formal process should not be anchored mainly on unsupported assertions or value judgments without serious efforts to establish objective proof and in this sense, it may be considered as part of "evaluative research" which utilizes systematic and "scientific research methods and techniques for the purpose of making an evaluation." Like evaluative research, it refers "to those procedures for collecting and analyzing data which increase the possibility for 'proving' rather than 'asserting' the worth of some social activity."[11]

For managers and for trainors, evaluation of training effectiveness presents conceptual as well as practical problems. Conceptually, the major problem lies in the difficulty of scientifically attributing to managerial training alone changes or improvement in both on-the-job performance or behavior at the individual level and

at the level of organizational performance because a number of factors and influences interact to affect individual or organizational performance. Determining effectiveness of training in on-the-job performance, in developing countries where resources for training are major constraints, has to be assayed in relation to other competing and compelling interests of the organization, particularly in the use of scarce financial resources.

Practical problems that confront decision makers in determining the need for evaluating training effectiveness in terms of performance in the work situation revolve around the question of which type of evaluation and measuring instrument could be used which will not entail enormous resources and where there is technical expertise within the organization. Will an evaluation which could predict future performance on the job (e.g., learning or immediate evaluation) suffice in the absence of capability (technical or financial resources) for the more rigorous types of evaluation? Finally, are there performance standards against which individual or organizational performance will be evaluated?

Evaluating the effectiveness of managerial training, unlike skills training for repetitive and mechanical tasks, is often more difficult because performance indicators cannot easily be measured.[12] Are training goals and objectives stated in measurable form so that qualitative or quantitative judgements could be made on whether these are attained, and what is the degree of attainment?

[11] Edward Suchman, *Evaluative Research Principles and Practice in Public Service and Social Action Programs* (New York: Russel Sage, Foundation, 1977), pp. 8-9.

[12] March G. Brown, "Evaluating Training via Multiple Baseline Designs," *Training and Development Journal,* July 1980, p. 11.

The Importance of Setting Measurable Objectives

It may be assumed that if the objectives of the management training is so general, for example, to improve the managers' attitudes, skills, and knowledge of the managerial environment, skills in various management tools and techniques, then it would be difficult to evaluate in a rigorously scientific manner the value of the training to the manager's job performance since this encompasses the broader field of his activity as a manager. However, some specific component of a general curriculum (e.g., use of specific management skill like project appraisal) could be measured in terms of both knowledge acquisition (learning evaluation) and application on the job (behavior or intermediate evaluation).

On the other hand, if the training objective is more concrete and specific, say, to improve his ability in financial analysis, or marketing, or project appraisal, then the task of measuring the changes in job performance will be less difficult. Thus, the more objectives are specific and clearly defined the greater the likelihood that they are measurable in quantitative or qualitative terms, for example, if one of the training objectives is to improve the managers leadership skills, these broad objectives must be broken down into the more specific dimensions of leadership which are measurable.[13]

The experience so far in developing countries seem to suggest a tendency to give a more generalized management training to executives who are drawn from various sectoral ministries, public enterprises, and in some cases, including private business, as in the case of the program of the Philippine Executive Academy. (See Table 1). The composition of the training group tends also to vary in terms of functional responsibility -- that is, they perform different task for their organization (finance, marketing, administration, etc.) and the only common denominator is their belonging to generally the same level in the managerial hierarchy. Thus, it is more difficult to frame more specific training objectives against which evaluation of training effectiveness will be measured. However, the adoption of the modular and block approaches to training managers, particularly for public enterprises, provides opportunities for evaluating training effectiveness using the basis of appraisal; for example, this approach in curriculum development which uses training modules flexibly in response to varying characteristics of the training group had been developed for training public enterprise managers.[14]

Evaluation Should Be Part of Course Design

Ideally, the evaluation of training effectiveness should be an integral part of planning, developing and im-

---

[13] Gabriel U. Iglesias, "Notes on the Evaluation of Training of Public Enterprise Managers," Regional Workshop on Management Training and Development in Public Enterprise," Regional Workshop on Management Training and | Developing in Public Enterprise, 10-15 January 1981, Karachi, p.6.

[14] See "Curricula for Training Public Enterprise Managers: Conceptual Framework and Approach to Curriculum Development for Training Rural Enterprise Managers" in Gabriel U. Iglesias, Sushil Chandra, and Melito Salazar (eds.), Training Public Enterprise Managers (Kuala Lumpur, Malaysia: Asia and Pacific Development Centre, 1980), pp. 34-62.

TABLE 1. FRAMEWORK FOR EVALUATING
(PHILIPPINE EXECUTIVE ACADEMY, 1969,

What is Being Measured	Who Are Evaluated	Other Evaluator
REACTION, EVALUATION Feelings on course and general experience with PEA o reactions after nominations o expectations from course o what he actually got o evaluation of course content o contribution of training staff o opinion of residential nature of course	Trainees	

LEARNING (INTERMEDIATE)
o self-rating on performance
   before and after course
o rating of participants'
   performance before and after
   course (both based on 16
   categories aimed at by the
   training)
	Trainees	Directing Staff

BEHAVIOR (INTERMEDIATE)
(Based on 16 categories)
o rating or participants'
   performances on the job
   (1) change (positive/negative)
      in the performance on
      the job
   (2) satisfaction with the perform-
      ance of the participants
      based on organizational
      standards both in terms of
      changes in:
      (a) use of specific manage-
         ment techniques
      (b) conceptual and
         analytical skills
      (c) general human
         relations skills
	Trainees	Supervisors

RESULTS (ULTIMATE) none		

Source: Philippine Executive Academy, *An Evaluation of Its Program* (Manila – 1969)

TRAINING EFFECTIVENESS: PHILIPPINES
EXECUTIVE DEVELOPMENT PROGRAM)

Timing	Evaluation Instrument	Analysis of Data	External/Internal Evaluator
after the course	Questionnaire and rating form	Average Rating and Tabulation	Within the training organization
	Interview question-naire with rating form	Average rating and Tabulation	
on the job months/years	Interview question-naire with rating form	Average rating and Tabulation	

1982

TABLE 2. FRAMEWORK FOR EVALUATING TRAINING EF
(Local Government Center, College of Public Adminis
Local Administration Develop

What is Being Measured	Evaluated/Evaluators	Timing of Evaluation
REACTION EVALUATION		
Reaction to the course (7 variables)	Participants	After the course
(1) achievement of workshop/ seminar objectives		
(2) general organization of the course		
(3) course content		
(4) assignment		
(5) time schedule		
(6) methodology		
(7) attainment of course objectives		
LEARNING EVALUATION		
Based on 6 factors	Directing Staff	During and
(1) participation in the classroom		Immediately after
(2) participation in discussion groups		the training course
(3) participation in group reports		
(4) preparation of individual reports		
(5) attitude		
(6) attendance		
BEHAVIOR EVALUATION		
(1) Favorable attitude changes	(1) Participants	Months up to
(2) Increased level of knowledge and skilled developed in various aspects of local government administration and development	(2) Superiors	more than two years
RESULTS EVALUATION		
None	None	None

Source: "The LADP Revisited and It's OK," *Local Government Bulletin*, Vol. XI, No. 1.

FECTIVENESS: PHILIPPINES
tration, University of the Philippines,
ment Program)

Evaluation Instrument	Analysis of Data	Internal or External Evaluation
Questionnaire	(no data)	Within the organization
Rating Form	Average Rating ("excellent", "very good", "good", "fair" "poor")	Same as above
Questionnaires, Interviews	Percentages	Same as above
None	None	None

1982

TABLE 3. FRAMEWORK FOR EVALUATING TRAINING

What is Being Measured	Who are Evaluated	Timing	Definition	Key Elements Being Measured
REACTION EVALUATION Feelings of trainees on contents, methods presentation, materials, training objectives, learning behavior[1]	trainees	(1) after every session (2) after training (3) after two years[8]	How well the trainees like the course	(1) Course content (2) Training methods (3) Presentation (4) Training objectives (5) Facilities (6) Training materials (7) Learning (8) Behavior[10]
LEARNING EVALUATION Amount of learning in terms of facts, principles and techniques[2] Knowledge, skills, attitudes and values	(1) trainees (2) trainees (experimental group) and control group	before and after course	What principles, facts and techniques were understood and absorbed by trainees	(1) Changes in attitudes (2) Changes in skills (3) Changes in knowledge
BEHAVIOR EVALUATION Changes on attitudes and behavior on the job[3] Changes leading to more effective job performance[4]	(1) trainees (2) trainees, superiors and subordinates (3) subordinates only[5] (4) trainees and subordinates[6] (5) trainees, peers, superiors, and subordinates	(1) on the job (2) before and after course[9]	(1) Application of principles and techniques learned in training on-the-job (2) Changes in behavior on the job and effect on job performance	(1) Changes in attitudes on the job (2) Changes in behavior on the job (3) Changes leading to more effective job performamnce[11]
RESULTS EVALUATION Change in unit/ organizational perform ance, such as reduction of costs, reduction of turnover and absenteeism, increase quality and quantity of production, and improved morale	(1) experimental (trainees) and control group[7] (2) trainees and supervisors	before course and on the job	Results classified as: Reduction of costs; Reduction of turnover and absenteeism; increase quality and quantity of production, improved morale	(1) Effect on performance of group/unit (2) Effect on performance of organization[12]

EFFECTIVENESS (UNITED STATES)

Importance	Evaluation Instrument Commonly Used	Analysis and Interpretation of Data
(1) Inputs to current and future training of favorable response conducive to learning (2) Measure feeiings and perceptions of (3) Measure subjective value of training	(1) Rating chart (for trainees) (2) Rating sheet (for Staff) (3) Standard evaluation form, Training material evaluation checklist	Tabulation and quantification

(1) Determine amount/ level of attitudes, knowledge and skills acquired
(2) Changes in course content, methods, etc.
	(1) Classroom Performance (Exercises)
	(2) Pencil and Paper Test (for measuring principles and facts)
	(3) Use of Standardized Tests
	  (a) "How to Supervise"[13]
	  (b) "Supervisory Inventory for Human Relations"[14]
	  (c) "Social Reaction Inventory of Human Relations"[15]
	(4) Use of questionnaire
		(1) Average rating
		(2) Correlational analysis
		(3) Multiple regression analysis[21]
		(4) Item analysis[22]

(1) Relate training effectiveness to job performance (2) Relate training to changes in behavior (application of ASK on the job)	(1) Questionnaires[16] (2) Attitude and opinion survey[17] (3) Attitude survey	Statistical analysis before and after
(1) Relate training to group, unit and organization performance (2) Relate training to solving management	(1) Records comparison (safety) (2) Depth interview questionnaire (cost reduction)[18] (3) Employee Relation Index (ERI) (extent of group of employees accepting and performing in accordance with the objectives and policies of the company)[19] (4) Likert study to measure periodic training[20]	

TABLE 3.   FRAMEWORK FOR EVALUATING

[1] Elton Reeves and I. Michael Jensen, "Effectiveness of Program Evaluation," in Donald Kirkpatrick (compiler), *Evaluating Training Programs* (Madison, Wisconsin: American Society for Training and Development (ASTD), 1975), pp. 42-46.

[2] Kirkpatrick, *op. cit.*, p. 6.

[3] Reeves and Jensen, *op. cit.*, p. 43, *passim*.

[4] Vera Kohn and Treadway Parker, "Some Guidelines for Evaluating Management Development Seminars," in Kirkpatrick, *op. cit.*, pp. 60-65.

[5] T.R. Lindholm, "Supervisory Training and Employee Attitudes," *ASTD Journal*, November-December 1953, as cited in Kirkpatrick, *ibid.*, p. 11.

[6] C.G. Moon and Theodore Hariton, "Evaluating An Appraisal and Feedback Training Program," *Personnel*, November-December 1958.

[7] P.V. Stroud, "Evaluating a Human Relations Training Program," *Personnel*, November-December 1959, as cited in Kirkpatrick, *op. cit.*, p. 12.

[8] Reeves and Jensen, *op. cit.*, p. 42.

[9] R.A. Fleishman, E.F. Harris, and H.E. Buntt, "Leadership and Supervision in Industry," (Ohio: Personnel Research Board, Ohio State University, 1955) as cited in Kirkpatrick, *op. cit.*, p. 11.

[10] Reeves and Jensen, *op. cit.*, pp. 42-46.

[11] Kohn and Parker, *op. cit.*, p. 60, *passim*.

[12] *Ibid.*, p. 60, *passim*.

[13] File and Remmers, as cited in Kirkpatrick, *op. cit.*, p. 7.

[14] Kirkpatrick and Planty, as cited in Kirkpatrick, *op. cit.*, p. 7.

[15] Julian B. Rotter, "Generalized Expectancies for Internal Versus External Control of Reinforcements," *Psychological Monographs*, 1966, 80, No. 1, pp. 1-28.

[16] Fleishman, Harris, and Buntt *op. cit.*, as cited in Kirkpatrick, *op. cit.*, p. 11.

[17] Floyd Mann, "Human Relations in the Industrial Setting," (Ann Arbor, Michigan: Survey Research Center, University of Michigan).

[18] Lester Tarnopol, "Evaluate Your Training Program," *ASTD Journal*, March-April.

[19] Developed by Willard Merrihue and Raymond Kartzell, *Harvard Business Review*, December 1955. See Kirkpatrick, *op. cit.*, p. 16.

[20] Rensis Likert, *Harvard Business Review*, April 1958. See Kirpatrick, *op. cit.*, p. 16.

[21] See Kohn and Parker, *op. cit.*, pp. 62-65, for use of average rating, correlational analysis, and multiple regression analysis.

[22] Joseph Gilbert, Henry Campbell, and Albert Oliver, "An Evaluation of Interdepartmental Training With Objective Tests," in Kirkpatrick, *op. cit.*, pp. 57-58.

TRAINING EFFECTIVENESS (FOOTNOTES)

Due to numerous footnotes in Table 3, the following footnote guide is presented by table heading:

what is being measured
      1, 2, 3, and 4
who are evaluated
      5, 6, and 7
timing
      8 and 9
definition
      none
key elements being measured
      10, 11, and 12
importance
      none
evaluation instrument commonly used
      13, 14, 15, 16, 17, 18, 19, and 20

analysis and interpretation of data
      21 and 22
reaction evaluation
      1, 8, and 10
learning evaluation
      2, 13, 14, 15, 21, and 22
behavior evaluation
      3, 4, 5, 6, 9, 11, 16, and 17
results evaluation
      7, 12, 18, 19, and 20

TABLE 4.   FRAMEWORK OF EVALUATION

	Type of Evaluation	What is Being Evaluated	Who are Evaluated	When Evaluated
GREAT BRITAIN	Reaction Evaluation	Trainees' reaction to training	Trainees	Final review sessions[2]
	Outcome Evaluation			
	oLevel 1—Immediate Outcomes	Changes in knowledge skills and attitudes	Trainees	After the course
	oLevel 2—Intermediate Outcomes	Effect of changes in knowledge, skills, and attitudes in promoting changes in job behavior	Trainees Managers of trainees and managers of control group/	Before and after training During course and six months later (for managers of trainees) After six months (for managers control group)
	oLevel 3—Ultimate	Effect of changes in on the job behavior of trainees leading to changes in departmental or organizational per-formance, in terms of output, costs, scrap rates, labor turnover, accident frequency, overall pro-fitability or effectiveness of the total organization	No data	No data

Source: * Peter Warr, Michael Bird, and Neil Rackham, *Evaluation of Management Training*, (London: Cower Press, 1971)

[1] Case N, p. 74.

[2] Cast C. p. 36.

[3] Case D, p. 42.

plementing a training program. But in developing countries where cost constraint, lack of technical expertise, and competing claims for attention and interests in the use of organizational resources are important decision factors, the more "systematic and rigorous" type of evaluation were generally done several years after the training institutions started their training programs as suggested in Table 1 on Philippine Executive Academy (PEA) and Table 2 on Local Government Center (LGC). Thus, there has been so far no attempt to evaluate training effectiveness conforming to the standards of evaluative research as far as training programs of developing countries included in this study. For example, one of the standard requirements of a rigorously scientific evaluation is the use of experimental (those who have undergone training) and control groups (a matched group who did not) to determine changes in behavior or performance on the job. (See Tables 3 and 4).

It is generally useful to incorporate training effectiveness evaluation in the planning and design stage of the course to ensure greater understanding and cooperation from participants and their organization about their purpose and the instruments to be used in the evaluation. Involving course participants in the design of evaluating instruments will not only increase its relevance and acceptance but also their cooperation when asked to complete the questionnaires or when they are subjected to in-depth interviews after they are back on their jobs. They will also be in a better position to clarify ambiguous items in the evaluation instrument to colleagues included in the evaluation sample (e.g., peers, superiors, and subordinates). This strategy will, hopefully, improve the generally low response rates of self-administered evaluation questionnaires.

The design of the evaluation instrument is extremely important since this would determine whether evaluation results are valid, reliable, and usable. If the training institution does not have qualified staff for this task, it either hires a consultant to assist in the evaluation or have the evaluation done by outside experts. In the long-term perspective, it will be less costly for the training organization to recruit qualified and experienced staff or to send them to training programs where they could study and learn evaluation of training.

Generally, better results can be achieved if evaluation of training effectiveness is done by the course participants and by their peers, superiors, and subordinates, since the participants' own evaluation, though useful, cannot provide reliable and comprehensive measures of training effectiveness. However, in view of the time and financial cost involved, it may be useful to explore other evaluation strategies.

Evaluation Strategy

There is clearly a need for and the desirability of evaluating training effectiveness to improve current and future programs in terms of finding out how much learning (whether knowledge, attitudes, and skills) have been transferred and acquired by the trainees. Equally important is to determine more accurately how learning made the difference in terms of improved performance on the job and by implication, improved per-

formance of the unit or the organization itself.

Evaluation of training should not be limited in measuring changes of learning and job performance. The evaluator/trainer should use the evaluation process as a mechanism in mapping out the future strategies which would asssit the trainees in applying the knowledge, skills and behavior acquired to the work situation. This could be achieved by preparing them to anticipate and to cope with potential re-entry problems in the organization.

This also means that the training organization would have developed a program based on a systematic assessment of training needs of the organization and have set specific and measurable training objectives. Once a decision to evaluate is made by the organization, alternatives are opened up in terms of types or techniques of evaluation to be conducted, what aspect of the training will be measured, which evaluation method or instrument will be used, who will be the objects of evaluation, when and how often will the evaluation be conducted, and how to use the findings of the evaluation as inputs to policy and organizational action.

The selection of alternative approaches and possible combination of approaches to and techniques in evaluating training effectiveness would depend largely on two important considerations: (1) the purpose or motivation behind the decision to evaluate training effectiveness; and (2) the organizational constraints, particularly resources, staff, technical competence, and so on.

Purpose and Constraints of Evaluating Training Effectiveness in Relation To Types of Evaluation Used

Although the experiences of developed and developing countries in evaluating training effectiveness found in Tables 1 to 4 may be considered a selective inventory, one could infer from the data not only differences in stages of development in general and training and its evaluation in particular (i.e., the "State-of-the-Art") but also the resource endowment and constraints for evaluation of training effectiveness. The developed countries' experiences (Tables 3 and 4), mainly drawn from the private business sector in America, reveal the following general characteristics and features:

(1) more extensive use of experimental and evaluative research which demand rigorous application of sophisticated research methodologies which are costly and time-consuming; for example, the use of experimental (those who had training) and control; (those who did not undergo training but with similar characteristics as the experimental) groups;

(2) the systematic and scientific use in descending order, of all four types of evaluation: reaction, learning (immediate), behavior (intermediate), and results (ultimate);

(3) the availability of standardized evaluation instruments; and

(4) the more frequent conduct of rigorous types of evaluation, particularly to measure changes in learning and changes in job performance.

These contrast sharply with the experi-

iences of developing countries, particularly those drawn mainly from Philippine experience. From the limited sample of evaluating training effectiveness among developing countries (mainly the Philippines), it would appear that the following types of evaluation were done: (1) reaction evaluation, expectedly, was commonly done either during the training, after each module, each session and after training; (2) learning evaluation (knowledge transfer) was conducted by the Philippine Executive Academy (PEA), the Local Government Center (LGC) and the Civil Service Academy (CSA); (3) behavior evaluation was done by the PEA, LGC, Development Academy of the Philippines (DAP); and, no results or ultimate evaluation was reported (See Tables 1, 2, & 5).[15]

The following observations may be drawn from the experiences of developing countries, especially the Philippines, on evaluating training effectiveness:

(1) Unlike the developed countries' experiences, there was no reported evaluation of training effectiveness in terms of results or ultimate evaluation; that is, measuring the impact of training on the performance at the individual or at the unit or organization level;

(2) The evaluation of learning (immediate) and evaluation of behavior (intermediate); that is, measuring changes in learning after the course and changes in the behavior and performance of the individual in on-the-job situation, did not use any experimental or control groups nor did they use before and after measures (including pre- and post-tests), except for the PEA (see table 3) where "before" and "after" evaluation was done through self-rating by the trainees as well as through rating by the training staff;

(3) There was heavy reliance on interview and survey questionnaires but there was no reported use of tests generally used by Americans to measure attitudinal and behavioral changes;

(4) Although three of the sampled training programs are long-term and recurring; that is, the Philippine Executive Academy (PEA) has been offering its courses regularly since 1966, the Local Government Center (LGC) since 1968, and the Development Academy of the Philippines (DAP) since 1973, only two (PEA and LGC)[16] have conducted a somewhat rigorous and systematic evaluation of training effectiveness by measuring the effects of training on behavior or performance on the job but the evaluation was done only once: the PEA in 1968 and the LGC in 1975. However, the DAP had initiated a Re-entry Project (REP) requirement as a method of determining learning transfer and applicability on the job.[17] Career Executive

---

[15] Unfortunately, studies from India, Pakistan, Bangladesh, and Indonesia are so sketchy on evaluation of training effectiveness to be included in this study. This represents a serious limitation in terms of comparative experiences.

[16] In the case of the Development Academy of the Philippines (DAP), lack of published material describing its evaluation prevented a more extended discussion of its evaluation methods and techniques.

[17] Starting with the 14th session in April 1976, the LGC's Local Administration and Development Program initiated a variation the DAP's re-entry projects. See *Work Programs*, 14th LADP, Local Government Center, College of Public Administration, University of the Philippines, 1976.

TABLE 5. FRAMEWORK FOR EVALUATING TRAINING EFFECTIVENESS:

Developing Countries	Type of Evaluation	Conducted by	What is Being Measured
PHILIPPINES Others: o Civil Service Academy (CSA) Junior Managers' Training (JMT)[1] o Development Academy of the Philippines (DAP) Career Executive Service Development Program (CESDP)[2]A/[2]B	REACTION EVALUATION	CSA, JMT DAP, CESDP	Reaction to training o impact on trainee of program o attainment of course objectives o over-all acceptability of the course Reaction of trainee on o program inputs.[2]A/[2]B
MALAYSIA o National Electricity Board (NEB), Management Development Programme for Middle Managers in Public Utility[3]		NEB	Reaction on the training in terms of o evaluation of lecturers, content, presentation o overall evaluation of th course
	LEARNING EVALUATION	CSA, JMT CESDP, DAP NEB	Change in Knowledge (No information) None
	BEHAVIOR EVALUATION	CSA, JMT NEB DAP, CESDP	None None Re-entry Project Approach to evaluate improved managerial capability back on the job[2]A/[2]B
	RESULTS EVALUATION	None for all	None

Sources: [1]Albina M. Dans, "The Junior Managers' Training," Regional Workshop on the Evaluation of Training Packages for Public Enterprise Managers, Bangkok, 10-14 August 1981. ICPE, APDC, Chulalongkorn.
[2]AAmelia Ancog, "The Career Executive Service Development Program," Human Resources Development, Manila: College of Public Administration, 25th Anniversary Conference, 9-11, June 1977.

DEVELOPING COUNTRIES

Who Are Evaluated/ Other Evaluators	Evaluation Instrument Used	Timing of Evaluation
Trainees	Questionnaire	After each module;
Participants/faculty	Evaluation format	after course
	Reaction (part- icipants)	During course-regular intervals After the course
	Baseline instruments	Before the course
Trainees	Questionnaire	After each session
	Questionnaire	After the course
	Verbal Evaluation	After the course
Trainees	Survey	
	Questionnaire	Before and after course
(No information)		
(None)		
Participants/ superiors	Re-entry project/ interview	Six months after course
None for all		

[2] BEmma Vineza, "Training Evaluation: DAP Experience," Report of Syndicate II, "Evaluation of Training," Regional Workshop on Management Training and Development in Public Enterprises of Developing Countries, Karachi, 10-15 January 1981, ICPE.

[3] National Electricity Board (LLN), *Report on Management Development Programme for Middle Managers in Public Enterprises* (Utility Sector), Bangi, Selangor: LLN Training Institute, April 29 to May 23, 1980.

Service in the Development Program (CESDP), and

(5) The evaluation of training effectiveness in the sample was generally conducted by in-house staff and there was no reported use of outside consultants to do the evaluations, with the possible exception of the PEA which had an American consultant working with the Filipino staff.

Types of Evaluation

Learning (Immediate) Evaluation

From the foregoing, it would appear that there is a need to strengthen the capability of training institutions in evaluating training effectiveness. For instance, there is a need to develop their ability to use the "before and after" measure, including the administration of pre-tests and post-tests. None of the sampled programs used this approach commonly employed in developing countries to evaluate changes in attitudes, skills and knowledge acquired during training (See Table 4) Lack of standardized tests could be one of the reasons. There are, however, many exercises which could be used to determine the ability of trainees in the use of certain concepts and skills (e.g., feasibility analysis, network analysis, team building and interpersonal skills, and leadership skills) through seat-work exercises, panel work, role playing, management games, simulation, and so on.

There is also a marked absence in the sampled training programs of the use of matched groups; that is, the experimental group (those who underwent the training) and the control group (those with similar characteristics but did not go through the training), a commonly used technique in advanced countries to increase the reliability and validity of findings.

While it is important to measure whether the training program actually transferred attitudes, skills, and knowledge to the trainees, there is no assurance that what has been learned is applicable and will be applied by the trainee on the job and if applied that the training led to changes in his performance and behavior.[18] However, learning or immediate evaluation is useful in providing data needed in improving future courses and to a certain extent, as a predictive instrument to future performance of the trainee in the work situation.

It may be inferred that the inability to use control groups and "before and after" measurements reflects both the lack of capability in technical as well as financial terms, thereby affecting the validity and reliability of existing evaluations. Apart from the above considerations, evaluative tests sometimes lead to positive unanticipated consequences; for example, studies have shown that the mere act of administering questionnaires and tests before the start of the training makes the trainee more receptive to the training material and to the learning itself.[19]

Furthermore, care should be exercised in developing the evaluation

---

[18] Virginia Scheir, "An Evaluation of A Long-Term Training Program," in Kirkpatrick, op. cit., p. 67.

[19] James Belascoa and Harrison Trice, "Unanticipated Returns of Training," in Kirkpatrick, op. cit., p. 67.

instrument and in determining the most suitable learning indicators to be measured, for example, participation in various aspects of the training process and attendance indicators used in the Local Government Center (LGC) study (Table 1) served as ambiguous and imprecise measures of learning. These indicators were used by the training staff to rate the learning performance of participants who were rated "excellent," "very good," "good," and "poor."[20] As Williams noted in his paper on the Local Administration Development Program (LADP) evaluation,

> The attempt to obtain some absolute measure of "excellence" on the basis of arithmetic average of scores is highly questionable. So is the result: the two lowest categories were empty showing that no participant was found to be anything less than "good."[21]

Behavioral or Intermediate Evaluation

As noted earlier, behavior or intermediate evaluation provides a mechanism which measures the effects of a training proramme or on-the-job performance or changes in the behavior."[22] The experience in the Philippines, based on a limited sample, seems to indicate that there is room for optimism in both the PEA and LGC evaluations. The PEA evaluation of training effectiveness in on-the-job situation appeared to be conceptually sound and employed a fairly sophisticated evaluation methodology in measuring behavior or performance on the job and in the interpretation of the findings (See Table 1).

The elaborate process by which the evaluators determined the training goals and objectives of the PEA, the preparation and the testing of the interview schedule and the rating forms, and the correlations done on the results of the evaluation showed a high degree of evaluative capability. It is unfortunate that this type of evaluation was done only once in 1968-1969 and was never repeated, although, the PEA has been offering regular courses from 1966 up to the present. The evaluation, which was administered on the supervisors of the trainee-executive, were also asked to rate the performance of executives who did not undergo the PEA training (control group). Thus, more valid comparisons could have been made and the evaluation could ascribe to the training changes in performance of the trainees matched with the control group.

The LGC evaluation, on the other hand, was administered to the trainees, his superiors, and his peers. It sought to measure changes in attitudes, skills and knowledge on the job. As in the PEA evaluation, no control group was included so there is no standard of comparison. The use of the multiple regression approach was suggested by one evaluator to reveal more meaningful correlations in the LGC study[23]

---

[20]Luz R. Oyales, "The LADP Revisited and Its OK," *Local Government Bulletin*, *Vol. XI, No. I* (January-June 1970) pp. 8-9.

[21]Arthur R. Williams, "The LADP Revisited — A Multiple Regression Approach to Evaluation," *Philippine Journal of Public Administration*, Vol. XXIII, No. 2, April, 1979, p. 214.

[22] See Schier, *op. cit.*, p. 109.

[23]Williams, *op.cit.*, p. 2.

The Re-entry Project Approach (REP) used by the DAP in evaluating the effectiveness of its training programme on the higher civil service deserves mention since it does not adhere closely to the standard approaches used in behaviour evaluation. To ensure that the participant will apply what has been learned during training, the participant/trainee will work out his Re-entry project plan with the training staff while still on training. After the training, the trainee will try to implement his re-entry project and the training staff evaluates his performance six months after graduation. The performance in implementing the REP will be one of the bases of graduation and for getting a Career Executive Service Officer (CESO) rank. The training staff not only evaluates the trainee but also interviews the superiors and subordinates of the trainee. The focus of the evaluation is not only the implementation of the REP but also the improved managerial capability back on the job.

Concluding Comments

Improving the capability of training effectiveness deserves attention since this study, although drawing on a fairly limited sample, can conclude that the "State of the Art" in effectiveness evaluation in developing countries showed weaknesses in the use of more rigorous and reliable methods. From the material appended, there is an evident lack of appreciation of the value of evaluation and failure to develop) and sustain — a strategy which will increase evaluative capability as well as sensitivity to and appreciation of the need to conduct more reliable and valid types of evaluation.

The use of REP and further improvements on the commonly used reaction evaluation instruments could yield less costly and reasonably useful evaluation of training effectiveness. Certain improvements could be done in the reaction type of evaluation so that it becomes a useful feedback instrument for future training programs. Efforts could be directed in developing standardized tests or in modifying existing ones to suit local conditions. There is also a need to professionalize the role of training evaluators within the organization and to incorporate evaluation in the development and planning process of training programs.