

## **Analysis of Evaluation Studies of Teacher Education Programmes in Punjab: Some Methodological Issues**

Khalid Mahmood\*

---

### **Abstract**

Punjab is the largest province of Pakistan with a population of more than 80 million— about 60% of the total population of the country. There are more than 63 thousand schools with approximately 325 thousand teachers. Following the educational reforms in Pakistan, the Government of Punjab has taken a number of new initiatives for teacher education (in-service and pre-service) in order to improve the quality of teachers' performance in government sector at primary, elementary and secondary levels. In order to evaluate these initiatives, a number of independent and/or sponsored studies were conducted. This paper has first developed a logical framework based on Guskey's model of professional development evaluation and then has critically analyzed the methodological issues of these studies using the framework. The analysis in the paper reveals that most of the studies were initiated with limited scope and hence lacked the comprehensive evaluation of these programs. None of the studies was a result of an inbuilt evaluation mechanism of any professional development programme. Some issues of these studies regarding timeline and instruments used for data collection, target population, and validation of data were also surfaced and highlighted through this paper. It is recommended that the evaluation mechanism should be an inbuilt part of the professional development programmes and should hold a formative nature.

**Keywords:** In-Service Teacher Education, Professional Development, Evaluation, Professional Development Evaluation Levels

---

### **Introduction**

The ultimate goal of a school is the overall development of students with more focus on their learning. The role of a teacher in this development is imperative. A large body of research advocates that the teacher remains the single most important school-based factor related to student learning (see Cooper & Jackson, 2005; Cooper, 2004; Darling-Hammond, 2000; Frampton et al., 2003; Levine & Cooper, 1991; Sanders & Rivers, 1996; Zohar, 2004). With a view to the significance of teacher factor for students learning it is recommended that continuous Professional Development (PD) of teachers is essential in order to keep them updated on the latest

---

\*IED, Aga Khan University, Karachi

developments to curriculum and effective in their selection and use of a repertoire of pedagogical skills. Nevertheless the primary focus of PD for teachers should remain the improvement of student learning (Collins 2000; Guskey 1999 & 2002b). Bissaker (2001) declares that “[teachers’ PD has] a significant influence on promoting positive learning outcomes for students” (p.41). Hence the purpose of teachers’ PD is to help teachers to become better teachers (Belzer, 2003; Elmore, 2002; Farnsworth et al., 2002; Guskey, 2002a; Killion, 2002; Lewis & Shaha, 2003).

Recently the PD of teachers has attracted policy makers as a focal area of intervention. Sykes (1999) points out that “professional development initially was regarded as one among a number of coequal policy instruments for promoting change, it is now reckoned as the centerpiece” (p.152). Similarly, Guskey (1986) also reminds us that “high quality staff [professional] development is a central component in nearly every proposal for improving education” (p.5).

As the importance of PD programmes is growing, evaluation of these programmes is also getting attention as an area for policy and research. However to evaluate the effectiveness of PD programmes and validating their impact, increasingly requires multidimensional approaches. Yates (1981) points out the difficulties in evaluating the effectiveness of PD programmes arising from the absence of a coherent evaluation strategy. However keeping in view the nature of PD programmes their evaluation can and should assess the effects of these programmes at multiple levels (see Belzer, 2003; Gorostiaga & Paulston, 1999; Guskey, 2002b; Killion, 2002; Kirkpatrick, 1998; Lewis & Shaha, 2003; Mizell, 2003; Smith, 1995).

This paper aims to analyze the research methodologies of the studies conducted to assess the impact of and/or to evaluate PD programmes in the province of Punjab, Pakistan. The paper specifically attempts first to develop a logical evaluation framework for PD programmes from available literature and then discusses the following questions in relation the developed framework:

- a. To what extent are the methodologies of the impact and/or evaluation studies conducted in Punjab relevant to evaluate PD programmes?
- b. What did these studies focus on in order to measure the effectiveness of the concerned PD programme?
- c. What are issues with the methodologies of these studies and how can we deal with these issues?

By addressing these questions, it is hoped that the study will; (a) contribute to a better understanding of PD programmes in general and their evaluation in particular and (b) help in planning and evaluating PD programmes.

## Context

This section provides a brief overview of the context in order to understand the breadth and the dynamics of the system for which the studies were conducted. There are four provinces of Pakistan — Balochistan, North Western Frontier Province (NWFP), Punjab, and Sindh —along with Northern and Federally Administered areas. Punjab is the largest province of Pakistan with a population of more than 80 million which accounts for almost 60% of the total population of the country. There are 35 districts in Punjab. In government sector, there are more than 63 thousand schools with about 325 thousand teachers and 9 million students. Among 63 thousand schools 44,175 are primary, 5,974 elementary (middle), 4,425 secondary, 304 higher secondary and 8,229 mosque schools (Govt. of Punjab, 2004).

There are a number of institutions in the government sector in Punjab which have been contributing in the PD activities of teachers. These include University of Education (UE), Institute of Education and Research (IER), Directorate of Staff Development (DSD), Provincial Institute of Teacher Education (PITE), two departments of education in universities, six University Colleges of Education (UCEs), and 36 Govt. Colleges for Elementary Teachers (GCETs).

In the past, following the educational reforms in Pakistan, the Government of Punjab took a number of initiatives for the PD of teachers through various programmes including in-service teacher training (Pardhan, et al., 2004; Saeed, 2002) to improve the quality of teachers' performance. Most of these initiatives focused on the PD of science and mathematics teachers and to some extent English teachers teaching at primary, elementary and secondary levels. The federal or provincial governments and foreign donors funded these programmes.

Keeping in view the spread of education in Punjab, almost every PD programme was designed to provide in-service training to thousands of teachers through different programmes in phases and rounds spanning over a period of two to three years. Further more the Cascade Model (Key or Leader Trainers ⇌ Master Trainers ⇌ Teachers) was used in imparting training to teachers under each project/programme. Teachers' learning was assessed through uniform pretests and posttests at all training centres throughout Punjab.

## Evaluation of PD

Despite the prevalence of PD activities and their numerous potential benefits, the effectiveness of these efforts cannot be ensured by good intentions alone. Evaluation is needed to measure and verify the programme quality and effectiveness. But in reality the evaluation of PD has been sparse and not many studies are available in the PD literature. The reasons for so

few evaluation studies could range from resource implications to a lack of evaluation skills to perceptions among educators and practitioners about the utility of evaluation studies. Guskey (2000) reports that often educators and other practitioners, view evaluation of PD programmes as waste of staff's valuable time and resources. Nevertheless evaluation is a critical part of the PD process because it identifies the programme elements which are (not) effective or successful, which ultimately leads to the creation of a more effective and efficient programme design.

Killion (2002) defines the term evaluation as “a systematic, purposeful process of studying, reviewing, and analyzing data gathered from multiple sources in order to make informed decisions about a programme” (p.42). However, the most common practice for evaluating the PD has been the satisfaction factor or as literature refers “happiness quotient”: What did you like the best? What did you like the least? What suggestions do you have for change? etc. Surely these are important questions, but they do not provide sufficient information to evaluate whether or not the PD brought about the desired changes in classroom teaching, in organizational support, or in students learning. During the last decade, several evaluation models have emerged which demand more than “happiness quotient” questions (Coffman, 1999; Guskey, 2002a; Killion, 1998; Kirkpatrick, 1998; McLaughlin & Jordan, 1999; Mizell, 2003). Guskey (2002a) believes that questions like did the topic address individual's needs? was her/his time well spent? or did s/he enjoy the refreshments? etc. are good to ask from a person who has attended the PD workshop in order to evaluate the workshop at its completion. But he contends that too often PD evaluation stops there and never even gets to the questions of effects or impact of these programmes. According to him “good evaluation can inform decision making about PD processes and effects” (2000, p. 2).

### *Developing a logical framework for PD evaluation*

PD experts agree that evaluations can and should assess the effects of the PD programmes at multiple levels— participants (teachers), classroom delivery, organization, students etc. (Belzer, 2003; Guskey, 2000 & 2002b; Killion, 1998 & 2002; Kirkpatrick, 1998; Lewis & Shaha, 2003; Mizell, 2003; Smith, 1995). Majority of these experts have identified levels of PD evaluation and they have placed various aspects of PD programme under these levels. These various aspects reflect the logical route of the programme. For example: IMSA (2001) emphasis on planning, implementation and outcome of the programme; Stufflebeam's (2002) famous CIPP model and McLaughlin and Jordan (2004) logic model view the route in terms of context, input, process and product. Some authors like Killion (2002) have given eight steps for such route.

Some of these models are very generic in nature and can be applied on any process for development e.g. CIPP model or logic model can equally be applied on any evaluation study, similarly Killion's (2002) eight steps encompass three phases (a) planning (b) conducting and (c) reporting and can also be followed in any evaluation study. Some of the models are particularly devised for the evaluation of PD programmes e.g. Kirkpatrick's (1998) four-level model and Guskey's (2002a) five-level model are more specifically to evaluate a PD programme.

However as compared to Kirkpatrick's four-level model that measures (a) reaction (b) learning (c) behaviour and (d) results, Guskey's (2002a) five-level model (see complete model in appendix-I) covers all requirements cited in Kirkpatrick's model except change in behaviour and attitudes of trained participants. Moreover, Guskey's model brought a new dimension in Kirkpatrick's model i.e. organization support and change. This aspect is necessary for creating a conducive environment for the growth or experimentation of any innovation or new learning.

Guskey has been writing for some time about the importance of seeking evidence of effectiveness in PD programmes for teachers and schools (see Guskey, 1985; Guskey, 1986; Guskey, 1990; Guskey, 1991; Guskey & Sparks, 1991; Guskey and Sparks, 1996; Guskey, 1998; Guskey, 1999; Guskey, 2000; Guskey, 2001; Guskey, 2002a; Guskey, 2002b). Furthermore his model also gives coverage to others models — covers almost all the four-level Kirkpatrick (1998) model; most components of CIPP model can be fit in within each level of evaluation; encompasses the steps in Killion (2002) model; containing more logical steps etc. The uniqueness of Guskey's (2002a) model is its ability to offers helpful way of thinking about gauging impact at different levels, and may be related directly to different orientations and intended outcomes. Another real advantage of this model could be seen in terms of its utility in planning of the PD programmes. The levels in this model could be kept in view while setting objectives of professional programmes and then implementation strategies could be devised on this basis. Guskey (2002a) describes the process for "working backwards" from "the student learning outcomes that one wants to achieve (Level 5)" and through each successive level to "what set of experiences will enable participants to acquire the needed knowledge and skills (Level I)."

For this study Guskey's model has been adapted and used with addition of another level for the analysis of the methodologies of impact/effectiveness studies on PD programmes.

#### *Five critical levels in Guskey's model of the PD evaluation*

Guskey (2002a) suggests five critical levels for evaluation of the PD programmes. The five levels in Guskey's model are hierarchical, ranging

from simple to complex. Following are the five levels with a brief description of each:

### *Participants' Reactions*

At the first level we gather participants' initial reaction about the PD activity / session / programme. A brief follow-up questionnaire — most of the time a standardized questionnaire — for participants is commonly used for this purpose. This questionnaire generally contains questions like: Did the participants think their time was well spent? Were the activities meaningful? Did participants think the activities would be useful in practice? etc. It is important to find out participant's reaction in order to improve the design and delivery of programmes.

### *Participants' Learning*

At the second level, evaluation measures participants' learning. It measures the knowledge, and skills that participants have acquired as a result of the PD activities, as well as the subjective indications of impact. This may involve more than a standardized questionnaire; for example, it may be necessary to use a paper and pencil exercise (e.g. test), a simulation or skill demonstration, oral or written personal reflections, portfolio evaluation, or similar activities. Guskey (2002a) warns against using merely a "standardized form" and instead urges "that the indicators of successful learning" should be designed to fit specific local needs. So whatever technique is used, the measures must reflect the goals for that activity, meeting specific criteria outlined before the PD experience begins. It will help in improving programme, content, format, and organization of the PD programmes. This level of assessment must also be structured to discover any unintended learning or results.

### *Organization Support and Change*

The third assessment level comes after an appropriate length of time has passed. This more complicated type of assessment analyzes organizational support for the skills gained in the PD programme. At this level we usually ask questions like: Was individual change encouraged and supported? Was administrative support public and overt? Were problems addressed quickly and efficiently? Were sufficient resources available, including time for sharing and reflection? Were successes recognized and shared? etc. It will lead to document and improve organizational support to inform future change efforts. This type of data could come through analysis of district education offices or school records or the minutes of follow-up meetings, questionnaires, or through structured interviews.

### *Participants' Use of New Knowledge and Skills*

The fourth critical level of evaluation assesses participants' use of new knowledge and skills. It can be assessed by asking whether teachers are using what they learned and using it well. It can rely on questionnaires or structured interviews, oral or written personal reflections, or examination of participants' journals or portfolios, but probably the best method is direct observation or observation via video or audio tape. Such measures should be ongoing with time intervals in between. It is important to document the utility of new knowledge and skills in actual classroom settings to improve the implementation of programme content.

### *Student Learning Outcomes*

At the fifth level, student-learning outcomes are measured. Guskey (2002a) acknowledges that it is impossible to determine simple cause-and-effect relationships between the PD programmes and improvement in student learning. He advises that in the absence of proof, we should collect good evidence, which does not have to be a lengthy and difficult process. At this level questions should be like: Did students show improvement in academic, behaviour, or other areas? Did the students benefit from the activity? Were there any unintended results? etc. It will guide to focus and improve all aspects of programme  
(For more detail about above-mentioned levels see appendix-I.)

## **Addition of a level in the Guskey' Model of the PD Evaluation**

### *Change in Teachers' Beliefs & Attitudes*

Guskey (1986) has identified a model of teacher change which advocates the same concept. According to that model,

“when teachers see that a new programme or innovation enhances the learning outcomes of students in their classes... then and perhaps only then, is significant change in their beliefs and attitudes likely to occur” (p.7).

Guskey (2002b) considers the PD programmes as a vehicle “to bring about change in classroom practices of teachers, in their attitudes and beliefs, and in the learning outcomes of students” (p.381). Moreover, he has identified two crucial factors contributing to the failure of the PD programmes: (1) teachers' motivation “to engage in the PD, and (2) the process by which change in teachers typically occurs” (p.382)

Keeping in view these failures and ultimate objective of the PD programmes i.e. positive change in students' learning and sustainability of

the positive change through change in teachers' beliefs and attitudes, we can add one more level in Guskey's Model i.e. Change in teachers' beliefs and attitudes.

When I discussed the addition of a level in the model with Guskey through mail, he responded:

"I discuss a model of change in teachers' attitudes and beliefs that I set forth in 1986 in the book on EVALUATING PD on pages 138-144. I do believe that such change is important and, as you point out, I contend that such change occurs only after teachers gain evidence of the impact of their efforts on student learning. In other words, their experience shapes their attitudes and beliefs. Hence, it seems to me that as a separate level, it would have to be added after level 5." (Guskey, 24<sup>th</sup> August 2005)

He gave a reason for not putting it as a separate level: "My hesitation to do so, however, was based on the idea that changing teachers' attitudes and beliefs is rarely considered a goal of the PD" (Guskey, 24<sup>th</sup> August 2005). However I would contend that change in teachers' attitude and beliefs is the ultimate goal of any PD programme.

Guskey (2002b) himself acknowledges that the "PD activities are frequently designed to initiate change in teachers' attitudes and beliefs and perceptions" (p.382). Kirkpatrick (1998) defines the behaviour change as the extent to which on-the-job performance has changed because the participants attended the PD programme. And according to Diem (2002) the impact of any programme is gauged from the ultimate change in people's attitudes or behaviour, or benefit in other ways. Wong (2004) also argues that moving forward to measure participants' change in behavior widens the focus of the PD evaluation studies. However, Guskey' model of the PD does not address this aspect of impact. This aspect becomes more important when it comes to sustainability in the utility of learning through the PD programmes.

Although the literature suggests that change in attitudes and beliefs is a gradual and difficult process for teachers (see Bolster, 1983; Fullan 1999; Fullan and Hargreaves, 1996; Guskey 1986; Lortie, 1975 etc.), its significance for the sustainability of the impact of the PD programmes cannot be overstated.

It is beyond doubt that we can use the "working backward" (Guskey, 2002a) concept for the planning purposes and for this we can still stay with five levels starting from setting standards for students learning outcomes. Nonetheless we have to target the change in teachers' beliefs and attitudes as a result of their PD for sustainability. Hence the ultimate goal of a PD programme should be change in teachers' attitudes and beliefs and from evaluation point of view we need to add this level in our critical levels

for evaluation.

Although it is challenging to observe the behaviour of a person over a long period of time in order to pass a judgment about change in beliefs and attitudes, there are methods that could be used to facilitate this. For example, we can get data through Retrospective Observations of their students, professional appraisal and/ or follow up visit with long intervals. Moreover, the challenging nature of the task should not hinder to bring sustainability in use of new learning in the classroom which the eventual purpose of the PD programmes.

### **Analysis of the methodology of the evaluation studies of PD programmes in Punjab**

After defining the framework of the PD evaluation, this part of the study will attempt to analyze the selected evaluation studies conducted in the province of Punjab.

### **Procedure and criteria for the analysis of the methodologies of PD evaluation studies**

About 20 studies are reported to have been conducted to evaluate the impact of the PD programmes in Punjab. Out of these, nine studies — Akhtar, 2005; Bhatti et al., 1998; Hussain, 1993; Johnson, 2000; Khan, et al., 2002; Mahmood, 1999; Nasir, 2000; Pardhan et al., 2004; and Saeed, 2001 — were accessible to me in the report form. These studies figured out not only the successes of these programmes but also pointed out the challenges that contributed to ‘poor’ PD opportunities. These studies have also put forward recommendations to improve the situation. These studies either were commissioned to external organizations by the Government of Punjab or initiated by the faculty of DSD or some other individuals out of their own interest.

Although the mode of delivery for these PD programmes was almost the same and they were also designed to cater for the same population, the studies to evaluate these PD programmes used different methodology. No matter what methodology a study uses, the important thing is what it really focuses on. To get the answer to this question six out of above-mentioned nine studies were selected as the sample. I have given representation to studies conducted for different reasons (purposes). Among these six; (a) two were commissioned to external organizations (b) two were initiated by the course organizing institutions (c) one was individual initiative of a course organizing institution’s faculty and (d) one was for M.Phil thesis.

The basic framework for the analysis of the methodologies was Guskey's (2002a) model with the addition of the sixth level. For this purpose, I focused on (a) main research question/objective to identify the focus of the study (b) duration of the training along with the period during which the training was conducted to see the validity of the data (c) population/sample to see the alignment of the data with study's focus (d) nature/type of the study to see the relevance of data and (e) instrument used by the studies and link them with various levels of Guskey's framework by analyzing nature of the questions in them. And (f) based on these factors decision was made about the level(s) touched by the study.

This paper does not aim to critique the studies but to see the patterns emerging from the studies for the PD evaluation.

## **Analysis of the studies**

### *Study i: Evaluation of Secondary Science INSET (Hussain, 1993)*

This study was conducted by an official of the Government of Punjab, Education Department. The aim of the study was to assess the impact of four-week in-service training courses of secondary school science teachers conducted in different rounds for different teachers during 1986 – 1990 at various pilot training centres established under Education Extension Centre (former nomenclature of the DSD) throughout Punjab. The objective of study was to:

“Get an insight into the organization, conduct and outcomes of in-service courses and to modify, update and if necessary completely change the organization and conduct of these courses, so that maximum efficiency [can be] achieved and proper utilization of man power and financial resource [can be] pursued.” (Hussain, 1993, p.6)

The study used a questionnaire comprising of both open and closed ended questions to meet this objective. Most of the items in the questionnaire were about facilities available at the training centres, adequacy of duration of the training, nature of activities, participants' learning especially on new content knowledge etc. the questionnaire was administered to over 110 training participants (Secondary School Teachers) who had attended the INSET. The data was collected 2 years after the closure of the programme.

It was a survey study. Although the focus of the study was to evaluate the programme nevertheless methodology of the study reflects that it was designed to solicit the training participants' reaction after 2 to 5 years of attending the training. According to Guskey's model participants' reaction is the first level of evaluation and it needs to be found out soon after

the closure of the programme.

The study also orally inquired about the participants learning. Nevertheless no formal instrument like paper and pencil test, participant reflections oral and /or written, or participants' portfolio was used to assess the acquisition of new knowledge and skills by them. The study tried to (a) measure participants' initial reactions with the experience and (b) assess their new knowledge and skills. In this way the study focused only on levels 1 and 2 of the developed framework.

*Study ii: Evaluation of present in-service training programmes for teachers and development of more cost effective method of INSET at various levels (Bhatti et al., 1998).*

This study was commissioned under Teacher Training Project (TTP) to evaluate the present in-service programmes conducted by the DSD and four foreign aided projects namely: Girl Primary Education Project (GPEP), Science Education Project (SEP) and TTP itself (funded by the Asian Development Bank) and Punjab Middle Schooling Project (PMSP) funded by the World Bank. The objectives of the study included assessment of the existing system of in-service teacher training and to recommend a more cost-effective system. The study was designed to answer the following questions specifically:

1. "What are the present systems of in-service education of teachers at primary, middle, high, higher secondary, and college levels?"
2. What is the coverage of content of these programmes under DSD and foreign aided projects at various levels?"
3. How effective are the courses organized at various levels?"
4. What could be a more cost-effective system?" (Bhatti et al., 1998, p.i)

It was a survey study. Separate questionnaire were used to collect data from trainers, teachers and headteachers of the training centres. A structured interview schedule for course organizers was also used. These instruments mostly contained questions about facilities available and management of the courses at the training centers. These instruments did ask participants about acquisition of the intended knowledge and skills and also gauged the understanding of the trainers on the knowledge and skills through obtaining their opinion. These instruments collected data from a sample of 1135 individuals including the training participants and trainers who had attended or conducted any of the above mentioned training programme(s) from 1992 to 1997. The data was collect from the participants who had

completed the training 1 to 5 years before the data collection.

The research questions and the instruments used for the study reflect that the scope of this study was broader than a PD evaluation study. The above-mentioned objectives 2 and 3 of the study probably were for evaluation. Keeping in view these objectives and the instruments used for the study, it is evident that the study tried to (a) measure participants' initial reaction to the experience, and (b) assess their new knowledge and skills. In this way, it could be said, the study came within reach of the level 2 of the framework. Instruments of the study also contained questions regarding procedures for nomination for the training, etc.

This situation reflects that the study asked questions for only 1<sup>st</sup> and 2<sup>nd</sup> levels of the framework, though time was not appropriate to ask such questions because the majority of the training participants may have forgotten the things happened at the training centers during the actual training.

*Study iii: Effectiveness of in-service training imparted through Teacher Training Project (Mahmood, 1999)*

This study was an independent study conducted by a student as a partial fulfillment of his M. Phil (Education) degree from Allama Iqbal Open University, Islamabad, Pakistan. The study tried to measure the effectiveness of in-service training imparted through Asian Development Bank assisted TTP (1993- 1999). The study was designed to answer questions regarding, (a) inputs provided in in-service training, (b) extent of change in competency of the teacher, (c) increase in students achievement level, (d) extent of change in behaviour of the teacher in the classroom and, (e) the implementation strategies adopted for teacher trainings (Mahmood, 1999). In order to get answers of these questions following instruments were used:

- Opinionaire for the teachers about the training course they had attended.
- Likert Type Five Point Scale to observe the attitudes of the teachers towards in-service training
- Achievement tests in the subject of mathematics and science for 5<sup>th</sup> grade students
- Retrospective Observation Schedule for the students of class 5 in order to know about their teachers' practices in classroom

It was an experimental study in which "The Static-Group Comparison Design" (Best & Kahn, 1998, p.148) was used. For this design control and experimental groups were identified at the time of data collection. Experimental group consisted of the teachers who had attended the training under the TTP from 1996 to 1997 and their students. 2400 students (1200 control and 1200 experimental) and 240 teachers (120 control

and 120 experimental) from two out of total eight regions of Punjab were included in the sample of the study. The data for the study was collected 1 year after the completion of the training (s).

Although this study addressed almost every level in the developed framework including sixth level, there were few shortcomings of the study. The initial reaction of the teachers and their assessment of the new learning through achievement tests was solicited one or two years after completion of their training, whereas Guskey (2002a) mentions data collection time for these levels is end of the training programme. This study did not probe properly about the organization support and change i.e. advocacy support, accommodation, facilitation, and recognition for level 3.

*Study iv: Evaluating the impact of in-service training: case studies of Punjab Middle Schools (Johnson, 2000).*

This study was one of the series of seven studies carried out under World Bank assisted PMSP (1993- 2000) on different issues. The study was conducted by three teams from different institutions. The study was funded by United Kingdom Department for International Development (DFID).

The study was designed to address the questions: "Is training received under MESH [Mathematics, English, Science, and Headteachers] evident in the professional practices of headteachers and teachers?" (Johnson, 2000, p.7)

The case study method was used to get the answer of this question. The case studies were conducted in 32 middle schools across Punjab. A total number of 70 teachers and 26 headteachers were included in these case studies. These teachers and headteachers had attended training programmes under the PMSP during 1995 to 1998. The data was collected one year after the completion of the training(s).

The focus of these case studies was to observe the utility of the training in the actual classroom setting and see its impact on the teachers' and students' behaviour and achievement, and on overall environment of the school.

This study touched level 5 of the framework. However, the study did not address the first level of the framework. Data for level 2 (orally enquiring about new learning) was collected after one to four years of completion of the training. Data needed for level 3 was collected through discussions with the teachers. Students' record (homework copies) was observed to assess the student learning outcomes.

*Study v: Impact evaluation of 5-day teaching skills development course for primary teachers in Punjab (Saeed, 2001)*

This study was an individual attempt of a faculty member of the DSD. The study aimed to assess the impact of the 5-day training conducted

during 1999-2000 in different rounds for different teachers under the DSD in two out of eight regions of Punjab. The study was designed to answer the question, “Is there any difference between the performance of trained and untrained PTC teachers in their classroom?” (Saeed, 2001). To answer this question pretest and posttest, classroom observation schedule and an achievement test for teachers, based on their understanding of the skills developed in the training were used.

Six thousand five hundred teachers had attended the training. The study used the same pretest and posttest developed by the DSD for the training and 5550 out of 6500 teachers’ pre and post test were included in the study. The classroom observations of 50 teachers (28 who had attended the training and 22 who had not attended the training from the same locality) were made to see the comparison. A separate achievement test was also administered to 156 teachers who had attended the training. Experimental design was used for the study and both “One-Group Pretest-Posttest” and “The Static Group Comparison” (Best & Kahn, 1998, pp. 147-148) were used. The data was collected immediately after the completion of the training.

The design of the study indicates that it was an effort to touch the level 4 of the framework but it did not capture levels 1 & 3. It did try to assess new knowledge and skills of the participants and its demonstration in the classroom. This was the only study, which used the pretest posttest design.

*Study vi: Effectiveness of in-service teacher education programmes offered by the University of Education, Lahore (Pardhan, et al., 2004)*

This study was a commissioned study funded by the DFID. The Aga Khan University – Institute for Educational Development and an NGO named “Society for the Advancement of Education” jointly conducted the study.

The study was designed to gauge the effectiveness of an in-service teacher training programme offered to 150,000 primary school teachers of science, mathematics and English. This training was conducted during 2001-2003 under the DSD and/or the UE. The study was conducted in 2003, during the last sessions of the training of English teachers. The study was undertaken to explore the following questions.

- a) “Was this the training the teachers needed?
- b) How well was the training designed and delivered?
- c) What was the impact of the training?” (Pardhan et al., 2004, p.10)

In order to respond to these questions, the study used individual / focus group interviews, questionnaires containing both open and closed ended items, classroom observation schedule, and document analysis

(instructional materials, programmes policy, planning documents and reports).

The sample of the study was selected from 9 different districts out of 35, across Punjab province. For focus group interviews 79 trainers and 149 teachers were included in the sample. For classroom observation two out of 9 already selected districts were chosen and 20 classrooms observations (10 in each of the two districts) were made. Thirty five training centres were visited which allowed observation of the delivery of the English language training programme. These observations were complemented by questionnaire, which was administered to 62 trainers.

It was the first study, which also investigated the basis of the training programme i.e. baseline data on which training was planned. The study tried to evaluate the programme at level 4 of the framework. The study collected data for initial reactions of the teachers either one year after the completion of the training or it was in progress. For level 2, study orally enquired about new learning and skills.

(For summary of the above analysis of the six studies see appendix-II.)

### **Common issues with the methodologies of evaluation studies**

Analysis of the above mentioned six studies brings out the following issues related to the impact assessment studies of PD programmes:

1. It is noticeable that none of these studies was a result of an inbuilt mechanism of evaluation of any PD programme. This phenomenon signifies an issue:
  - i) *Non-involvement of the training participants at designing stage of the study*
2. All six studies were conceptualized after the completion of their respective PD programmes which made it difficult for these studies to (a) capture the initial reactions of training participants and (b) gauge their new learning at the end of the training. Guskey (2002a) recommends suitable time for these two aspects to be at “the completion of the PD programme” (p.47).

Almost every study tried to enquire about the participants’ “initial” reaction after the passage of certain amount of time—ranging from a year to five years. These studies gauged participants learning through their verbal or written expressions about what they learnt from the PD programme. The most common form of measuring cognitive outcomes is through testing. Standardized and non-standardized testing forms a key part of the educational system, and is usually considered to provide the most reliable measure of cognitive outcomes (Muijs and Reynolds 2002 & 2004). So validity of the data for these two levels remains questionable. This situation reflects two important issues:

- ii) *Improper timeline for data collection*
  - iii) *Use of inadequate instruments for data collection*
3. Only less than half of these studies addressed a vital level i.e. organization support and change which has a major influence on the participants' use of new knowledge and skills. As Guskey (2002a) has mentioned, "lack of organization support and change can sabotage any PD effort, even when all individual aspects of PD are done right" (p.47). This highlights one more issue:
- iv) *Ignoring data that have significant effect on the process for bringing effect to PD programmes*
4. In almost all the studies data collection was a single snapshot. However, Guskey's model demands data collection especially regarding the use of new knowledge and skills i.e. level 4 as ongoing process with regular intervals of time. This situation indicates another issue with these studies.
- v) *Insufficient evidences to prove what has happened*
5. Only two out of six studies tried to measure the impact on students' learning. This brings up another issue.
- vi) *Ignoring target population or the ultimate beneficiaries in the sample of the study.*

## Discussion

Majority of the evaluation studies in Punjab ignored level 3 of the developed framework by ignoring data collection on organization support and change. In order to get the full advantage of the PD programme importance of studying organization support and change especially for large scale the PD activities becomes imperative. As Nasir (2000) found while looking for impact of the PMSP training in Punjab,

"As regarding feasibility of applying PMSP methodology, for quite a large number of teaches it was quite impossible due to many local environmental constraints" (p.32).

These constrain included large classes, non-supportive attitude of the headteachers and other supervisory staff, assessment system, unwelcoming behaviour for innovations etc. The two studies, which have addressed this level, had solicited teachers' opinion on this aspect instead of analyzing school records or records of other administrative office as suggested in Guskey's model. Palumbo (in Lincoln, 1985) has referred to this characteristic as "to be effective, evaluations need to be based on realistic assumptions about the nature of organizations and the implementation process" (p.13). It is further supported by Rocheleau (in Rogers and Hough, 1995) while discussing the need to incorporate administrative and organizational factors into evaluation designs though a

significant factor “Evaluation researchers have curiously failed to include organizational or administrative variables in their research designs” (pp.321). Hence PD evaluation studies should not ignore looking into organization support and change.

Moreover, without having a clue about participants’ learning, each of these studies except one tried to gauge course participants’ use of new knowledge and skills in actual classroom setting, through one way or another.

Furthermore, most of the studies (i.e. 4 out of 6) have not tried to judge the students learning outcomes as a result of their teachers’ PD. Loucks-Horstey and Matsumoto (1999) have mentioned that many researches regarding the impact of the PD have examined changes in the instructional practices, teachers’ knowledge, teachers’ belief, and other important variables that may be indirectly linked to students’ achievement [but not the students’ achievement itself]. There might be so many reasons behind it but Huffman et al. (2003) pointed out, “research on the impact of PD on students’ achievement is limited because it is difficult and expensive to study, and the link between the PD and students’ achievement is complex” (p.378). Guskey (2002a) has also supported this argument that

“The relationship between the PD and improvement in students’ learning in the real-world setting is far too complex and includes too many intervening variables to permit simple causal influence” (p.50).

Nevertheless he did not discourage evaluating change in students’ learning after their teachers’ PD. He has further mentioned that “but in the absence of proof, you can collect good evidence about whether a PD programme has contributed to specific gains in student learning” (p.50). In the case of Punjab, only two out of six studies attempted to gauge students learning. Out of these two only one has tried to gauge cognition of the students.

As a conclusion, analysis of the methodologies of the studies reveals that there is a possibility that different studies had focused different aspects of the PD programmes. Further, perhaps it was not possible (or for some experts may not be advisable) to study impact at every level through a single study. It might be due to lack of expertise in this field. As Guskey (2002a) points out, there are three major mistakes generally committed in the evaluation of the PD. First, evaluation may amount to not more than documentation of activities completed over a period of time. Secondly the evaluation may not go deep enough, and third, evaluation may be too brief. Just as PD should be an ongoing process, so should be its evaluation, no matter, whether these programmes are on a small scale or otherwise.

## Recommendations

In the light of the above discussion regarding the various methodological issues of the PD evaluation studies, it is recommended that:

- Methodology of an impact/evaluation study for a PD programme should also be planned while developing the programme. In other words every PD programme, especially large scale programmes should have an inbuilt evaluation study.
- The PD evaluation studies should be longitudinal and aligned with the timelines of the PD activities and their objectives.
- An impact study should not fail to notice factors affecting the impact of a programme especially, which are not linear with PD activities e.g. organization change and support.
- Appropriate identification of population strata for data collection is curial for every study especially PD evaluation studies. In this regard we should not ignore students who are the ultimate beneficiary of any development in the field of education.
- Instrument(s) should be developed in accordance with the kind of data required for each level of PD programmes. Especially to measure cognitive abilities, tests, portfolio, piece of work etc. should be taken into account.

## References

- Akhtar, Z. (2005). *Impact of in-service training on elementary school teachers in the subject of mathematics during 2002*. Unpublished M. Phil thesis submitted to Department of Secondary Teacher Education, Faculty of Education, Allama Iqbal Open University, Islamabad.
- Belzer, A. (2003). Toward broadening the definition of impact in PD for ABE practitioners. *Adult Basic Education*, 13(1), 44-59.
- Best, J. W. & Kahn, J.V. (1998). *Research in Education* (8<sup>th</sup> Edition). Boston: Allyn & Bacon.
- Bhatti M. .A. et al.,(1998). *Evaluation of present in-service training programme for teachers and development of more cost effective method of INSET at various levels*. A study report by the Pakistan Education Foundation, an NGO and submitted to Punjab Teacher Training Project, Directorate of Staff Development, Punjab Education Department Lahore.

- Bissaker, K. (2001). Did they get what they came for? Evaluating teachers' learning. *International Education Journal*, 2(4), 41-52.
- Bolster, A. S. (1983). Toward a more effective model of research on teaching. *Harvard Educational Review*, 5(3) 294 – 308.
- Coffman, J. (1999). Learning from logic models: An example of a family/school partnership program. Cambridge, MA: Harvard Family Research Project.
- Collins, D. (2000). *Achieving your vision of PD: How to assess Your Needs and Get What You Want. (3rd Ed.)*. North Carolina: SERVE, Greensboro.
- Cooper, E. J. (2004). The pursuit of equity and excellence in educational opportunity. In D. Lapp, C. Collins-Block, E.J. Cooper, J. Flood, N. Roser, & V. Tinajero (Eds.), *Teaching all the children: Strategies for developing literacy in an urban setting* (pp. 12–30). New York: Guilford.
- Cooper, E. J., & Jackson, Y. (2005). *The Importance of PD to Unlock the Potential of Students in Urban Settings: A Response to Virginia Richardson and Patricia Anders*. In J. Flood, & P. L. Anders (Eds.), *Literacy Development of Students in Urban Schools* (pp. 231-240). Newark, DE: International Reading Association
- Darling-Hammond, L.(2000). Teacher quality and student achievement: A review of state policy evidence. *Educational Policy Analysis Archives*, 8(1). Retrieved December 24, 2005, from <http://epaa.asu.edu/epaa/v8n1>
- Diem, K. G. (2002). *Measuring impact of educational programmes, a fact sheet*. New Jersey: The State University of New Jersey
- Elmore, R. F. (2002). *Bridging the gap between standards and achievement: The imperative for PD in education*. Washington DC: Albert Shanker Institute.

- Farnsworth, B. J., Shaha, S.H., Bahr, D. L., Lewis, V. K., & Benson, L. F. (2002). Preparing tomorrow's teachers to use technology: Learning and attitudinal impacts on elementary students. *Journal of Instructional Psychology*, 29(3), 121-138.
- Fennema, E., Carpenter, T. P., Franke, M. L., Levi, L., Jacobs, V. R., and Empson, S. B. (1996). A Longitudinal Study of Learning to Use Children's Thinking in Mathematics Instruction. *Journal for Research in Mathematics Education*, 27(4), 403 – 434.
- Frampton, P.; Vaughn, V.L. & Didelot, M.J. (2003). The PD school partnership: Is practice improving? Teachers and principals respond. *Journal of Educational Administration*. 41(3), 292-309
- Fullan, M. G. & Hargreaves, A. (1996). *What's Worth Fighting for your school?* New York: Teachers College Press.
- Fullan, M. G. (1999). *Change Forces: the sequel*. Bristol: Falmer Press
- Gorostiaga, J. M. & Paulston, R.G. (1999). Mapping new approaches in program evaluation; a cross culture perspective. ERIC Document Reproduction Service No. 428 433.
- Govt. of Punjab (2004). *Ready Reckoner on Punjab Education Statistics & EMIS achievements update 30-06-2004*. Lahore: Punjab EMIS Centre.
- Guskey, T. R. (1998). The age of our accountability. *Journal of Staff Development*, 19(4), 1-9.
- Guskey, T. R. (2000). *Evaluating PD*. Thousand Oaks, CA: Corwin Press.
- Guskey, T. R. (2001). JSD Forum: The backward approach. *Journal of Staff Development*, 22(3), 60-72
- Guskey, T. R. (1985). *Implementing mastery learning*. Belmont, CA: Wadsworth.
- Guskey, T. R. (1991). Enhancing the effectiveness of Professional Development programmes. *Journal of Educational and Psychological Consultation*, 2(3), 239-247.

- Guskey, T. R. (1990). Cooperative mastery learning strategies. *Elementary School Journal*, 91(1), 33-42
- Guskey, T. R. & Sparks, D. (1996). Exploring the Relationship Between Staff Development and Improvements in Student Learning. *Journal of Staff Development*, 17(4), 73-76.
- Guskey, T. R. (1986). Staff development and the process of teacher change. *Educational Researcher*, 15(5), 4-12.
- Guskey, T. R. (1999). Apply time with wisdom. *Journal of Staff Development*, 20(2), 10-15.
- Guskey, T. R. (2002a). Does it make a difference? Evaluating PD. *Educational Leadership*, 59(6), 45-51.
- Guskey, T. R. (2002b). PD and teacher change. *Teacher and Teaching: theory and practice*, 8(3/4), 381-391.
- Guskey, T. R. (2005). *Re: Addition in critical levels of PD evaluation: a response of an email from "Thomas R. Guskey" To: [ikhaldm@yahoo.com](mailto:ikhaldm@yahoo.com) Date: Tue, 23 Aug 2005 22:05:57 -040.*
- Guskey, T.R. & Sparks, D. (1991). What to consider when evaluating staff development. *Educational Leadership*, 49(3), 73-76.
- Harris, A. (2002). *School Improvement: what's in it for schools?* London: Routledge Falmer Taylor & Francis Group.
- Heibert, J. (1999). Relationships Between Research and the NCTM Standards. *Journal for Research in Mathematics Education*, 30(1), 3-19
- Huffman, D., Thomas, K. & Lawrenz, F. (2003). Relationship between PD, teachers' instructional practices, and the achievement of students in science and mathematics. *School Science and Mathematics*, 103(8) 378- 387.
- Hussain, M.M. (1993). *Evaluation of secondary science INSET*. A report submitted to the Education Extension Centre, Punjab Education Department, Lahore.

- IMSA [Illinois Mathematics and Science Academy] (2001). *Assessing the Impact of Educator PD Programmes*. Aurora: Office of Educational Research and Evaluation.
- Johnson D. (Ed) (2000). *The impact of in-service training: case studies of Punjab Middle Schools*. Lahore: Punjab Middle Schooling Project, New Muslim Town.
- Khan, H., Hafeezullah., Ahmad, T., Shah, D. & Amin, N. (2002). *Impact evaluation of the secondary school science education project in Pakistan (Research study No. 166)*. Islamabad: Academy of Educational Planning and Management.
- Killion, J. (1998). Scaling the elusive summit. *Journal of Staff Development*, 19(4), 12–16.
- Killion, J. (2002). *Assessing impact: Evaluating staff development*. Oxford, OH: National Staff Development Council.
- Kirkpatrick, D. L. (1998). *Evaluating Training Programmes: The Four Levels* (2<sup>nd</sup> Ed). San Francisco: Berrett-Koehler Publishers.
- Levine, D.L., & Cooper, E.J. (1991). The change process and its implications in teaching thinking. In L. Idol & B. Jones (Eds.), *Educational values and cognitive instruction: Implications for reform* (pp. 387–410). Hillsdale, NJ: Erlbaum.
- Lewis, V.K. & Shaha S.H. (2003). Maximizing learning and attitudinal gains through integrated curricula. *Education 123* (3).
- Lincoln, Y. S. (Ed.) (1985). *Organizational theory and inquiry: the paradigm revolution*. Beverly Hills, CA: Sage
- Lortie, D.C. (1975). *Schoolteacher: a sociological study*. Chicago: IL University of Chicago Press
- Loucks-Horstey, S, & Matsumoto, C. (1999). Research on PD for teachers of mathematics and science. The state of the scene. *School Science and Mathematics*, 99 (5) 258- 271.

- Mahmood, K. (1999). *Effectiveness of in-service training imparted through Teacher Training Project*. Unpublished M. Phil thesis submitted to Department of Teacher Education, Faculty of Education, Allama Iqbal Open University, Islamabad.
- McLaughlin, J. A. & Jordan, G. B. (2004). Using logic models. In Wholey, J., Hatry, H. P., & Newcomer, K. E. (Eds.). *Handbook of practical program evaluation* (2nd ed., pp.7-32). San Francisco: Jossey-Bass.
- McLaughlin, J.A. & Jordan, G.B. (1999). Logic models: a tool for telling your program's performance story. *Evaluation and Program Planning*, 22(1), 65-72.
- Mizell, H. (2003). Facilitator: 10, Refreshments: 8, Evaluation: 0. *Journal of Staff Development*, 24(4), 10–13. Available at [www.nsd.org/library/publications/jsd/mizell244.cfm](http://www.nsd.org/library/publications/jsd/mizell244.cfm).]
- Muijjs, D. & Reynolds, D. (2004) *Effective Teaching* , London; Sage.
- Muijjs, D. & Reynolds, D. (2002). Teachers' beliefs and behaviors: What really matters? *Journal of Classroom Interaction*, 37(2), 3-15.
- Nasir, A. (2000). *The Impact of In-service Training: Case Studies of Punjab Middle Schools*. A report submitted to the Directorate of Staff Development, Punjab Education Department, Lahore and Punjab Middle Schooling Project, Lahore.
- Pardhan, H. et al., (2004). *Effectiveness of in-service teacher education programmes offered by the University of Education, Lahore*. A study report by the Aga Khan University (AKU-IED), Karachi, and Society for the Advancement of Education (SAHE) an NGO, Lahore and submitted to the DFID and the Education Department Government of Punjab, Lahore.
- Rogers P.J. and Hough, G. (1995). Improving the effectiveness of Evaluations: making the link to organizational theory. *Evaluation and Programme Planning*, 18 (4), 321-331.

- Saeed, M (2001). *Impact evaluation of 5-day teaching skills development course for primary school teachers in Punjab*. A report submitted to Directorate of Staff Development, Punjab Education Department, Lahore.
- Saeed, M (2002). Impact evaluation of 5-day teaching skills development course for primary school teachers in Punjab. *Journal of Elementary Education*, 12 (1-2), 36-44.
- Sanders, W.L., & Rivers, J.C. (1996). *Cumulative and residual effects of teaching on future students: Value-added research and assessment*. Knoxville: University of Tennessee.
- Smith, H. (1995). *Rethinking America: Innovative strategies and partnerships in business and education*. New York, NY: Random House.
- Stufflebeam, D.L. (2002). *CIPP Evaluation Model Check List: a tool for applying the fifth installment of the CIPP model to assess long-term enterprises*. [Available at <http://www.wmich.edu/evalctr/checklists/cippchecklist.htm>]
- Sykes, G (1999). Teachers and student learning: strengthening their connection. In Darling –Hammond, L and Sykes, G (Editor) *Teaching as the learning profession: Handbook of policy and practices*. San Francisco: Jossey-bass.
- Wong, P (2004). The PD of school principals: insight from evaluating a programme in Hong Kong. *School Leadership and Management*, 24(2), 139-162.
- Yates, M. (1981) *Practice issues in the provision of opportunities for professional development of secondary heads*. Wakefield: City of Wakefield Education Department.
- Zohar, A. (2004). *Higher order thinking in science classrooms: students' learning and teachers' PD*. Dordrecht: Kluwer Academic Press

**Guskey’s Model of Five Levels of PD Evaluation (2002a)**

<b>Evaluation Level</b>	<b>What Questions Are Addressed?</b>	<b>How Will Information Be Gathered?</b>	<b>What is Measured or Assessed?</b>	<b>How Will Information Be Used?</b>
1. Participants’ Reactions	Did they like it? Was their time well spent? Did the material make sense? Will it be useful? Was the leader knowledgeable and helpful? Were the refreshments fresh and tasty? Was the room at right temperature? Were the chairs comfortable?	Questionnaires administered at the end of the session.	Initial satisfaction with the experience	To improve programme design and delivery
2. Participants’ Learning	Did participants acquire the intended knowledge and skills?	Paper-and-pencil instruments simulations Demonstrations Participant reflections (oral and/or written) participant portfolios	New knowledge and skills of participants	To improve programme content; format, and organization
3. Organization support & change	Was implementation advocated, facilitated, supported? Was the support public and overt? Were problems addressed quickly and	District and school records minutes from follow-up meetings Questionnaires structured interviews with	The Organization’s advocacy support, accommodation, facilitation, and	To document and improve organization support to inform future change efforts

	<p>efficiently?  Were sufficient resources made available?  Were successes recognized and shared?  What was the impact on the organization?  Did it affect the organization's Climate and procedures?</p>	<p>participants and district or school administrators  participant portfolios</p>	<p>recognition</p>	
<p>4.  Participants' use of new knowledge and skills</p>	<p>Did participants effectively apply the new knowledge and skills?</p>	<p>Questionnaires structured interviews with participants and their supervisors participant reflections (oral and/or written) participant portfolios direct observations video or audio tapes</p>	<p>Degree and quality of implementation</p>	<p>To document and improve the implementation of programme content</p>
<p>5.  Student learning Outcomes</p>	<p>What was the impact on students?  Did it affect student performance or achievement?  Did I influence students' physical or emotional well-being?  Are students more confident as learners?  Is student attendance improving?  Are dropouts decreasing?</p>	<p>Student records school records Questionnaires Structured Interviews with students, parents, teachers, and/or administrators participant portfolios</p>	<p>Student learning outcome cognitive (Performance and Achievement)  Affective (Attitudes and Dispositions)  Psychomotor (skills and Behaviours)</p>	<p>To focus and improve all aspects of programme design, implementation, and follow-up to demonstrate the overall impact of PD</p>

Summary of Analysis of Methodologies of the Six Studies

Levels of Evaluation	Study 1	Study 2	Study 3	Study 4	Study 5	Study 6
	Evaluation of Secondary Science NSET (1993)	Evaluation of present in-service training programmes for teachers and development of more cost effective method of INSET at various levels (1998)	Effectiveness of in-service training imported through Teacher Training Project (1999)	Evaluating the impact of in-service training: case studies of Punjab Middle Schools (2000)	Impact Evaluation of 5-day teaching skills development course for primary teachers in Punjab (2001)	Effectiveness of in-service teacher education programmes offered by the University of Education, Lahore (2004)
When data were collected	after 2 to 5 years of completion of the training (s)	after 1 to 5 years of completion of the training (s)	after 1 to 2 years of completion of the training (s)	after 1 to 3 years of completion of the training (s)	after 6 months of completion of the training (s)	after 1 year of completion of or during the training
1. Participants' Reactions	✓ using questionnaire to assess initial satisfaction	✓ using questionnaire to assess initial satisfaction	✓ using questionnaire to assess initial satisfaction	×	×	✓ focus group interview to assess initial satisfaction
2. Participants'	✓ using oral	✓ using oral reflection	✓ using	✓ using oral	✓ using pre-post	✓ using oral

<b>Learning</b>	reflection to assess new knowledge and skills of participants	and written opinion to assess new knowledge and skills of participants	achievement tests to assess new knowledge and skills of participants	reflection to assess new knowledge and skills of participants	tests and achievement test to assess new knowledge and skills of participants	reflection to assess new knowledge and skills of participants
3. <b>Organization support &amp; change</b>	x	x	x	✓ enquiring verbally from the teachers to assertion the department's advocacy support, accommodation, facilitation, and recognition	x	✓ enquiring verbally from the teaches to assertion the department's advocacy support, accommodation, facilitation, and recognition
4. <b>Participants' use of new knowledge and skills</b>	x		✓ using retrospective observation to measure the degree and quality of implementation	✓ classroom observation to measure the degree and quality of implementation	✓ classroom observation to measure the degree and quality of implementation	✓ classroom observation to measure the degree and quality of implementation

<b>5. Student learning Outcomes</b>	x	x	✓ using achievement tests to measure student learning outcome about all the three aspects (cognitive, attitudes and skills)	✓ analyzing students' piece of works to measure student learning outcome about all the three aspects (cognitive, attitudes and skills)	x	x
<b>6. Change in Teachers' Beliefs &amp; Attitudes</b>	x	x	✓ using retrospective observation to assess the change teachers' belief and attitudes	x	x	x

Key      ✓ represents data collection at a particular level  
x indicates no data collection on this level (aspect)