Bergmann, Herbert

Monitoring in multilevel teaching and learning situations

ZEP : Zeitschrift für internationale Bildungsforschung und Entwicklungspädagogik 24 (2001) 1, S. 22-28

Zeitschrift für internationale Bildungsforschung und Entwicklungspädagogik

24. Jahrgang März 1 2001 ISSN 1434-4688D

Impressum

ZEP - Zeitschrift für internationale Bildungsforschung und Entwicklungspädagogik 24. Jg. 2001, Heft 1

Herausgeber: Gesellschaft für intelkulturelle Bildungsforschung und Entwicklungspädagogik e.V.
Schriftleitung: Annette Scheunpflug
Redaktionsanschrift: ZEP-Redaktion, Professur für Bildungsforschung, Karl-Glöckner-Str. 21 b, 35394 Gießen
Verlag: Verlag für Interkulturelle Kommunikation (IKO), Postfach 90 04 21, 60444 Frankfurt/ Main, Tel.: 069/784808; ISSN 1434-4688 D

Erscheinungsweise und Bezugsbedingungen: erscheint vierteljährlich: Jahresabonnement DM 36.- Einzelheft DM 9,50; alle Preise verstehen sich zuzüglich Versandkosten; zu beziehen durch alle Buchhandlungen oder direkt vom Verlag. Abbestellungen spätestens acht Wochen vor Ablauf des Jahres.

Diese Publikation ist gefördert vom Ausschuß für Entwicklungsbezogene Bildung und Publizistik, Stuttgart. Das Heft ist auf umweltfreundlichem chlorfreien Papier gedruckt.

Herbert Bergmann

Monitoring In Multilevel Teaching And Learning Situations

Zusammenfassung: Evaluation sollte auch in der Entwicklungszusammenarbeit, z.B. im Projektmanagement im Bildungssektor eine wichtige Rolle spielen. Der Autor zeigt mit dem Modell des "Monitoring " eine Möglichkeit auf, die sich besonders in der Implementation von Innovationen im Bildungswesen bewährt hat.

The context of monitoring

The term has become prominent in development co-operation. The concept of M&E - Monitoring and Evaluation - is stock in trade in development planning and project / programme implementation. It became an issue in the late seventies and has become ever more important since then.

The emphasis on M&E is motivated by the wish to know whether a development project is on the right track or not - does it achieve its objectives, and if so, how well does it do so? The shift towards neoliberalism in the West in the early eighties, with its criticism of the role of the state, brought along an emphasis on accountability of public spending. Since development co-operation was mostly financed by taxpayers' money and since it was not very popular anyway, accounting for the *proper* and *effective* use of resources became a must. This would have called for periodic evaluations and audits. But in order to achieve good results, short-term monitoring had to precede evaluations. Only then could one hope to correct mistakes before an evaluation exercise would show major deviations from plans or unwanted effects that could only be corrected at high costs.

Monitoring and multilevel teaching/learning situations

What exactly is monitoring? Monitoring is a management tool for short-term optimisation of an ongoing work programme. It is the systematic, frequent process observation of the implementation of a project at predefined stages, using a limited set of indicators. It aims to detect weak points, errors, and unwanted effects early enough to correct them before major damage is done. As such, it is particularly important in situations that carry major risks, where experience is lacking, and where successful, effective routines are not yet in place. Multilevel teaching/learning situations fall into this category. There are different types of monitoring such as input monitoring, process monitoring, output monitoring and impact monitoring. *Input monitoring* focuses on the timely delivery of pre-specified inputs in the required quantity and quality. *Process monitoring* checks the timing and quality of ongoing work processes. *Output monitoring* looks at whether the immediate output of a work programme is reached (did the participants of a training course learn what they were supposed to). *Impact monitoring* looks at the effects of such programmes on the final beneficiaries. The monitoring approach presented in this paper is a mixture of process, outcome and impact monitoring.

Monitoring is in itself a process with a number of steps such as developing a monitoring design, data collection and analysis, drawing conclusions, and taking corrective action where appropriate.

Multilevel teaching/learning is a major option when large numbers of people have to learn something new in order to

Table 1: Multilevel teaching/learning situations in INSET, Pakistan and Indonesia

	Pakistan (PEP-ILE 2000)	Indonesia (SEQIP 2000)
Central Level: Project Team	n = 4	n = 4
Level 1: Master Trainers/ Local Consultants	n = 46	n = 40
Ratio Central : Level 1	1 : 12	1 : 10
Level 2: Learning Coordinators/ Advisory Teachers	1.550	2.640
Ratio Level 1 : Level 2	1 : 34	1 : 66
Level 3: Teachers/ Science Teachers	27.971	31.680
Ratio Level 2 : Level 3	1 : 18	1 : 12
Beneficiaries	962.445	1.604.679
Pupil:Teacher Ratio	1: 34	1: 51

PEP-ILE Primary Education Programme - Improvement of Learning Environment
SEQIP Science Education Quality Improvement Project

apply it in a short time so that direct face-to-face teaching by a small number of instructors is not feasible. The other option is distance education, which has a number of limitations when it comes to practical skills.

Multilevel teaching/learning situations form a cascade. Examples are (1) country- or province-wide teacher in-service training (INSET), (2) Adult education, e.g. literacy programmes with large numbers of participants, and (3) In-service training of professional staff in any sector employing large numbers of staff, e.g. health or rural development. The approach always consists of a cascade of direct face-to-face instruction where the trainees at one level are the trainers at the next level down. At the topmost level, there is a professional team designing curriculum elements (content, teaching / learning materials and teaching methods) that should promote better learning (in our case 4 persons). At the lowest level, there are pupils as the ultimate beneficiaries of the effort. In between, there are levels of multipliers (see table 1).

Table 2: Quality Loss: Test Results at Three Levels of a Cascade in Pakistan, Grade 3 End-of-Course Assessment, Per Cent of Maximum Score

	Group	N	Total	Mathematics	Language
Level 1/ Master Trainer	Total	74	82,17%	77,25%	85,68%
Level 2/ Learning Co-ordinator	Male	819	70,10%	68,51%	71,23%
	Female	359	71,22%	65,67%	75,19%
	Total	1178	70,44%	67,65%	72,43%
Level 3/ Teacher	Male	2089	64,19%	64,32%	64,10%
	Female	1157	65,97%	62,71%	68,30%
	Total	3246	64,83%	63,75%	65,60%

Source: PEP-ILE Monitoring Records

Basic Issues

A number of issues make multilevel teaching/learning a risky and difficult approach that needs close monitoring. These issues are (1) quality loss as one moves down the cascade, and (2) the nature of innovations, since most often, the message transmitted in the cascade is an innovation in the respective context.

Quality Loss

Any teaching/learning cascade with n intermediate levels has n+1 transition points where the curriculum is passed on from one level to the next. Since at each level, the delivery method is face-to-face instruction in small groups, learners at each level have to master the curriculum in such a way that they can teach it error-free at the next level. Experience shows that this assumption is not realistic. Transition points in a cascade are critical points for monitoring. Data from Pakistan illustrate the quality loss from Master Trainers to teachers (see table 2).

The comparison across cascade levels becomes possible because the same INSET programme is taught at each level. Therefore, the same test can be applied. After three years of continuous training, Master Trainers still don't master all they should, even at the level of recall. The quality loss between Master Trainers (Level 1) and Learning Co-ordinators (Level 2) is about 12 percentage points, between Master Trainers and Teachers (Level 3) about 17.

Introducing Pedagogical Innovations

These risks are stronger when the cascade system is used to introduce innovations on a large scale. Innovations are core elements of development co-operation. The main aim of co-operation is to improve the way things are done in a sustainable manner, and this involves innovations most of the time. New equipment, new methods and procedures are not necessarily innovations in any absolute sense, they might already be known and used for quite some time elsewhere. But they are new in the country where they are being introduced by means of development co-operation.

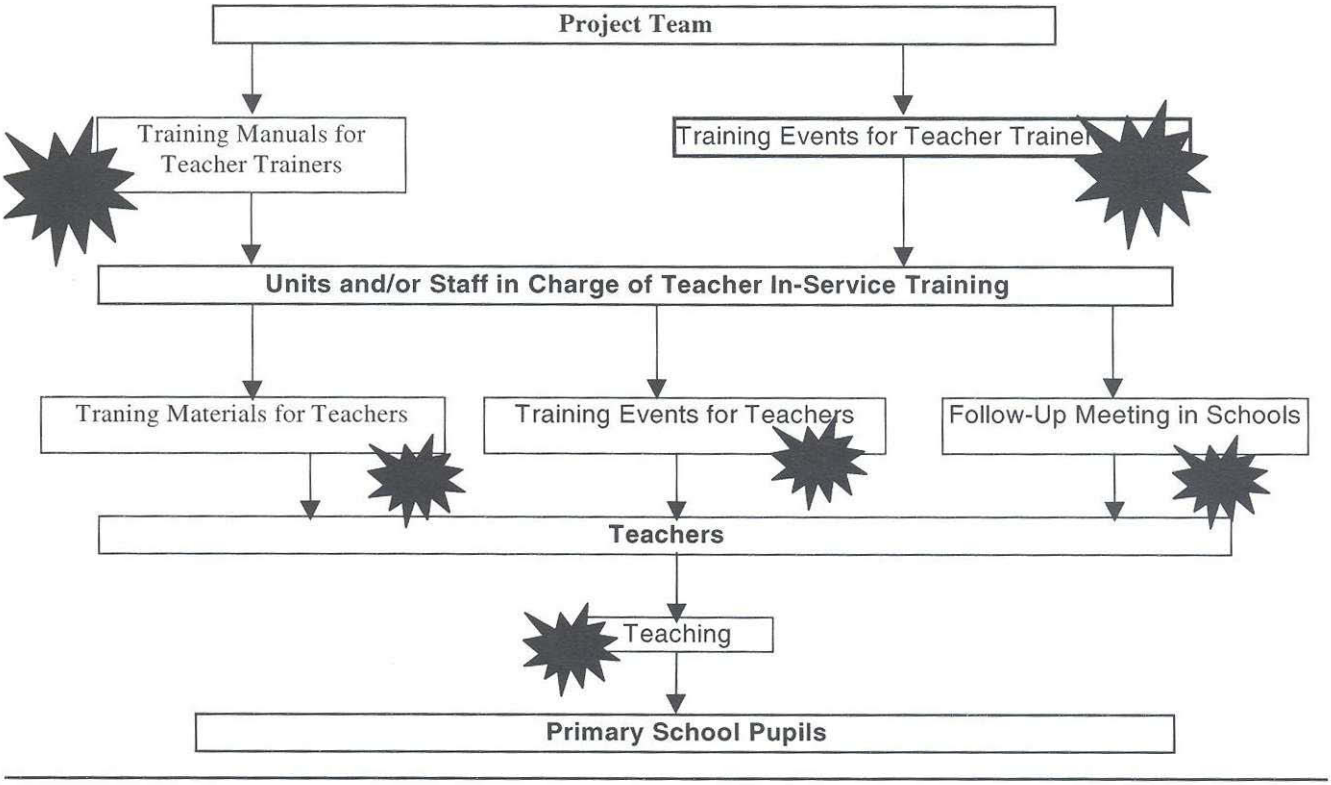
Examples are new content (environmental education, practical subjects, health education), new approaches to Basic Education (optimising the language of instruction by switching to mother tongue teaching), new teaching methods (the child-centred approach to teaching and learning, e.g. in science education), or new methods of school management (school development, school based management).

There are three major problems with innovations that make INSET programmes to disseminate them particularly difficult.

1. Because they are new, there is a lack of experience of how they work beyond protected settings like a model school. And therefore, the instructors might not be fully familiar with them, might not be in a position to explain and demonstrate them convincingly. The course programme might not address certain important issues. This needs to be monitored and will lead to course revisions.

2. They ask teachers to change their way of doing things. As long as "only" new content is at stake, this might not raise too much resistance. Certain teachers might simply not be able to understand it and would need remedial action (coaching, additional documents, an additional course).

Table 3: Critical Points in the Cascade



Changes in teaching methods, in school management, or in the language of instruction are prone to induce much more resistance. Beyond the technical difficulties of mastering new ways of doing things, they can touch upon matters of social status and professional identity: teaching in a formerly despised language goes against established value systems; changing towards a "pupil-active" classroom interaction can be resented as an attempt to undermine the teacher's unquestioned authority, etc. It would be wrong to ignore these aspects in the course programme, and therefore, they have to be identified, and tackled as required.

3. An innovation might have design flaws that become apparent only during implementation in the usual, everyday school situation. These need to be detected as soon as possible.

Monitoring as a method to manage the quality and the feasibility risk

The objectives of monitoring in multilevel teaching/ learning situations are to detect risks in order to limit and reduce them, to identify and solve problems before they cause extensive damage, and to improve things "on the go", i.e. during implementation. These objectives relate monitoring closely to action research. In fact, if there were sufficient time to conduct large-scale try-outs, cascade monitoring could focus on transmission errors at the critical points.

In order to reach these objectives in time, cascade monitoring focuses on four key elements, (1) the objectives of the educational project, (2) the critical points in the teaching / learning process, (3) the minimum data requirements, and (4) participation across all the levels of the cascade. Such an orientation is necessary to produce the right results at the right time.

Focus on Objectives: The objectives defined during project planning need to be reached and provide the main scope of monitoring. At each level of the cascade, there are two objectives: getting course participants to learn up to mastery (output objective) and satisfactory performance of the instructors (process quality objective). The latter objective is instrumental in reaching the output objective, and it must be monitored in order to be able to act if the learning results at the next level are not up to expectations.

Focus on Critical Points: Not everything in the process of multilevel teaching/learning will be monitored. For maximum efficiency, only critical points should be included. Critical points are all those elements and steps where the contents of the teaching/learning process are passed on to a different group of people. The materials and the events that contain the messages are all critical points. To illustrate this, the figure below presents a situation in Peru with only one intermediary between the project team and the classroom teachers. The INSET curriculum and all its supporting documents are developed by the project team. The first critical point is the transfer of the course content from the project team to those who are in charge of the INSET activities with teachers. Two elements are relevant, the training manual

developed for the INSET instructors, and the training events organised for them. Both need to be monitored. The training manuals need to be subjected to a formative evaluation, e.g. by a panel of professionally recognised external experts. The Training-of-Trainers courses need to be monitored like all training events. The next transfer occurs between the INSET instructors and the teachers. Instructors run INSET courses, distribute training materials for the teachers, and visit them in their schools to observe them teaching and advise them. Finally, teachers who have benefited from INSET courses teach their pupils.

All these elements constitute critical points. Errors in any one of them will be multiplied at the next level. Given the modest level of teacher training and pedagogical competencies of INSET staff in many countries, errors presented at a given level will hardly be detected by course participants. The higher up in the cascade an error is committed, the higher its potential for multiplication across the whole system, without much chances of onward correction by participants. It is therefore of utmost importance to include the uppermost levels in monitoring (see table 3).

Focus on the Essentials: Parsimonious data collection is a must. Only the minimum data required to reach the objectives of the monitoring must be collected. This is a major difference with research, and social science research in particular, where there is a tendency to collect more data than strictly needed by the research hypotheses because this would permit to test additional hypotheses that are generated during data analysis. Focus on the essentials also means leaving out all those variables that might explain differences in outcome (learning results of course participants) and process quality (performance of the instructors) but that cannot be acted upon (age, gender, ethnic, regional and religious affiliation, etc.). This is necessary to obtain monitoring results at the required time: results that come in too late to make corrections are of no use.

Focus on Participation: Cascade monitoring needs to be participatory right across all the levels of the cascade. One reason is that monitoring is often confused with control, and induces fear. This in turn defeats the very purpose of monitoring because, among other things, it leads to data "doctoring". In order to avoid this, those in charge of monitoring at any level need to know that it is not the message of monitoring results (yes/no, the programme was/was not successful) but the quality of monitoring (reliable, useful feedback) that will count towards their own evaluation.

Monitoring needs to associate all those who participate in the management of an innovation. In large-scale INSET programmes, centralised monitoring goes against its very purpose: errors can turn up at any level anywhere in the geographical area covered, and therefore, they need to be detected there and then. Responsibility is shared across levels and districts, and remedial action has to be undertaken as soon as possible. All this requires the full participation of everyone in charge.

The objective of monitoring has two important implications: management requires the capacity to act fast as soon as weak points and shortcomings are detected. If after data collection, data analysis, and reporting, there are no more resources to take corrective action, then the whole exercise does not make much sense. This means that this type of monitoring, like any other one, has to be a routine element of the overall management process, not an add-on. And it means that contingency resources have to be set aside for remedial action (coaching, additional reading material and documents, additional courses, visits, as the case may be).

Monitoring as a management tool

Requirements: As a management tool, monitoring needs to comply with the requirements of rapidity, relevance, and cost effectiveness. As a tool for short-term optimisation of an ongoing work programme, it needs to produce its results fast. Data and results that are not immediately relevant to what one is trying to achieve have no place here. Results

Table 4:

Level	Actors	Indicator Set I	Indicator Set II
Project Unit	Local and international expert teams	Quality of suggested intervention (innovation)	Performance as trainers
Intermediate I	Master Trainers	Learning results (degree of mastery)	Performance as trainers
Intermediate II	INSET instructors	Learning results (degree of mastery)	Performance as trainers
Intermediate III	Teachers	INSET learning results (degree of mastery)	Performance as teachers
Beneficiaries	Pupils	– Baseline data – Learning results (degree of mastery)	

coming in too late to correct weaknesses in an ongoing programme are next to useless.

Tools: Tools for this kind of monitoring are appropriate indicator sets, instruments for collecting data, and methods of data analysis. For long-term purposes, database tools are also required.

Indicators: It would be tedious to go into any detail concerning indicator sets. We shall present the structure of such a set that would allow all those in charge to assure the quality of the multilevel process (see table 4).

Indicators and instruments have to be developed according to the respective situation (subjects covered, type of innovation). Not all indicators will be needed at the same time. The need for baseline data in the area covered by the intervention is obvious: they are needed to establish the effect of the intervention on the ultimate beneficiaries. In addition, the intervention cannot be properly designed unless the baseline is known. One might say that all indicators other than pupils' learning results are unnecessary. As long as results improve as planned, this is right. But if this is not the case, one needs to know the cause and its location. This implies knowledge of the causal path used by the intervention.

Instruments for Data Collection: The usual instruments used for data collection in social science research are being used. These are test forms concerning the course content, and observation sheets for classroom observation. The course writers develop simple test forms as part of their assignment.

In order to push the analysis further, e.g. in order to find out why new content is not taught or teaching methods presented during INSET are not practised, in-depth interviews would be used. These could be individual interviews. Often, however, it is more useful to conduct group interviews, e.g. with course participants right at the end of the course, after the course evaluation data have been analysed, or after visiting teachers in a school cluster.

Methods of Data Analysis: Data Analysis is orientated towards decision-making. There is room neither for time-consuming exploration nor for sophisticated analytical techniques, e.g. in the sense of causal analysis. Scoring techniques, the use of percentages, presenting results by relevant subgroups (boys and girls, teachers by gender, ethnic group or mother tongue) becomes important.

Use in Decision-Making

Monitoring results are action orientated. They point to areas where the training programmes offered do not produce the intended results or where a shift of emphasis is required. Well-analysed and well-presented monitoring results lead to the following decisions:

To continue a successful programme: This is the case when the overall degree of mastery of the courses offered at various level is satisfactory as shown by the mean scores; the overall standard deviations indicate that there is no sizeable number of participants who remain below an acceptable standard. Results by subgroups of the course participants show that no particular group has been "left behind". Observations show the same when participants apply what they have learnt.

No further data analysis to be done: If everything works out to satisfaction, raw data and results are simply stored for further reference.

To change the programme: This is required when the overall degree of mastery of the courses offered at various

Table 5:Average Test Results by Gender and Participation in the Programme

	Total			Male			Female		
Score	0 yrs	I year	2 yrs	0 yrs	I year	2 yrs	0 yrs	I year	2 yrs
Pakki									
Urdu	22.5	28.8	38.2	26.3	35.0	38.4	14.1	19.5	38.1
Maths	34.1	44.3	57.2	40.2	50.9	59.5	20.7	34.4	53.8
Grade 2									
Urdu	18.4	23.8	35.1	18.7	26.3	30.8	17.7	20.1	41.5
Maths	32.1	40.4	56.0	36.5	46.6.	60.9	22.6	31.2	49.9
N	335	350	540	230	210	320	105	140	220

Source: Pakistan, PEP-ILE Impact Study 2000, unpublished manuscript

Notes: Schools are organised by gender, this concerns pupils and teachers alike. "Year" refers to the number of years schools have participated in the PEP-ILE.
Pakki is the second year of Grade one.

level is unsatisfactory - overall average results are too low, the overall standard deviations show that a large number of participants must have very low levels of mastery and are not fit to serve as instructors at lower levels or as teachers. Observation shows that very many have not developed the intended pedagogical skills. They cannot apply what they were presented during the courses.

Further data analysis is required: In cases such as those mentioned above, more in-depth analysis of the whole group of course participants at each level is needed.

- Sometimes, this might mean examining the distribution of results in detail. Is it normal or does it exhibit several peaks, pointing to subgroups with different performance levels?

- It certainly means examining every item, in both test forms and observation sheets: Are there content points or areas that are particularly difficult for the participants, or is there unsatisfactory performance across the board? Depending on the outcome, different lines of action will be taken.

- It also means examining results by subgroups - gender, educational background, levels of experience, etc.

- In many cases, it might mean additional data. Expert panels or group interviews with representatives of course participants would be asked to look into difficulties, elucidate the most probable causes, and suggest ways of coping.

Scope and type of remedial measures in the presence of an overall weakness: if the analysis shows that the weakness discovered through monitoring is general, i.e. neither restricted to certain content and skill areas nor limited to identifiable categories of participants, then basically two lines of action are open:

- To rethink the programme in its entirety, complete with objectives and measures, from scratch. Results suggest that the project team has severely misjudged the situation.

- If the performance level is not too far below the minimum requirements, then the courses need to be redesigned, starting at a lower level and allowing more time to build up foundation knowledge and skills. This decision needs expert judgement based on a lot of experience in similar situations.

Scope and type of remedial measures in the presence of a weakness limited to parts of the curriculum: if the analysis shows that the weakness discovered through monitoring is restricted to certain content and skill areas and applies more or less to all categories of participants - men, women, participants with differing educational background, different levels of experience, etc., then three lines of action are open:

- To adjusted the whole programme for "next time around", reinforcing the difficult areas,

- To develop an ad-hoc programme focusing on the weak areas and to recall participants with particularly low scores, offering them the ad-hoc programme as a remedial measure,

- To develop reading material in the respective areas and distribute it to all the course participants

Scope and type of remedial measures in the presence of a weakness limited to identifiable groups: If the analysis shows that the weakness discovered through monitoring is limited to identifiable categories of participants, then the following lines of action are open:

- To assess whether these groups need to remain in the programme. This could be an option at the level of instructors, but never at the level of teachers.

- To assess the difficulties the identified groups have with the course programme in more detail. Qualitative methods would be of help, based on an item analysis done for the identified groups alone - what exactly is difficult for this group of people, and why is this so?

- Once this is done, to develop an ad-hoc programme focusing on the weak points of the groups in question and to recall participants with low scores belonging to the group, offering them the ad-hoc programme as a remedial measure,

- To develop reading material in the respective areas and distribute it to all course participants of the group.

Resource Provision: All this needs resources that have to be planned beforehand. Since monitoring results are not foreseeable, these resources would fall under contingency planning. So far, there is little experience as to how much needs to be allocated for this purpose. It might be safe, for a start, to put it at 10% - 20% of the costs of the INSET programme at each level.

Monitoring Levels

As we have seen above, not all monitoring activities related to one INSET programme occur at once, and some might not occur at all. It all depends on the results of the first-level monitoring. The first level monitoring looks at all the events in a cascade - "all" meaning either every single course and single follow-up visit or a representative sample - which approach is chosen depends of the size of a programme. The approach could be labelled "quick and dirty", using a limited set of simple indicators in simple instruments. It just must not be too quick or too dirty to overlook important weaknesses. Again, deciding on this is a matter of professional judgement. No clear-cut rule can be formulated.

Once the weak points have been spotted with the help of first-level monitoring, the next level of monitoring would be "activated", an in-depth analysis of problem areas. As has already been said, mostly qualitative methods of data collection and analysis would be used.

Process quality monitoring and impact monitoring

The type of monitoring suggested here includes pupils as the ultimate beneficiaries in any education project. Since sustainable learning results of various kinds are the only short-term impacts of basic education, monitoring in multilevel teaching/learning situations combines process and outcome monitoring at all levels with impact monitoring at the beneficiary level. A last table from Pakistan will illustrate this point (see table 5).

Comparing the zero-year control group with the schools participating one and two years in the programme shows a cumulative effect on pupils' learning. This can be considered an impact, particularly so if the impact study planned for 2001 confirms these findings. The example also shows how data sets generated through monitoring lead into evaluation: if in a time series, no impact is visible, this raises doubts as to whether the intended objectives can be achieved under the prevailing circumstances. This is a question typically asked at evaluation. It elucidates another point: if monitoring activities are to be useful for later evaluation, they need to build a systematic database allowing analysis over time, across geographical / administrative units and social categories.

Monitoring and research

Concerning its pragmatic objectives and its nature as a management tool on the one hand, many of its data collection and analysis methods on the other, monitoring looks like a sophisticated instance of action research.

No opportunity for scientific research of the classical type?

With all the streamlining towards short-term quality management, impact assessment, and limited comparisons over time, do the data collected for monitoring have any use in scientific research? The author thinks they do.

1. In order to assess the effectiveness of the intervention against other key factors, these need to be included among the variables. Thus, the most important of the variables usually considered in research on learning results will finally be included in the data collection instruments. Otherwise, the relevance of one's intervention and the sustainability of the effects cannot be properly evaluated.

2. Given the multiple level of intervention (pupils get books, teachers are trained, classrooms are improved, school management and the district administration are improved), data sets need to be structured so as to make multilevel analysis possible. Otherwise, the effects of these various levels on learning cannot be properly assessed.

3. Tests and observational data can be subjected to techniques such as factor analysis in order to look for underlying dimensions.

All data collected has to be reliable and representative, if it is to guide action. Given the above considerations, this type of monitoring will therefore generate data that can be analysed beyond their immediate use. In most developing countries, there are hardly data of the level of quality produced by monitoring as described in this paper. They permit quite a range of secondary and comparative analyses.

The relevance of monitoring in more developed countries

Despite the rapid advance of ICT (Information and Communication Technology) and the promises it holds for learning and course delivery, multilevel teaching/learning situations will continue to be used for INSET programmes. Although the potential for self-learning and peer learning at school level is higher than in most developing countries, and distance delivery modes are much more feasible, certain topics will probably be presented through a multilevel approach. It would be useful to monitor these programmes in the same way that development co-operation programmes are monitored. One reason is that in our countries, too, accountability concerning public spending has become an issue.

Remark:
The author would like to thank his colleagues and friends, both national and expatriate, in basic education projects sponsored by GTZ in Indonesia. Pakistan and Peru for the valuable insights he could gain while working with them on monitoring systems and systems of impact evaluation.

Dr. sc.pol. Herbert Bergmann ist augenblicklich Seniorfachplaner für Grundbildungsvorhaben bei der GTZ und zuständig für Projekte in Asien. Er promovierte in Soziologie und ist seit 1974 in der Entwicklungszusammenarbeit tätig. Erfahrung in Projekten sammelte er in Kamerun, Peru und Ruanda. Von 1995 bis 1997 war er in der Weltbank als Planer von Bildungskrediten tätig und arbeitete mit den Ländern Gabun, Kamerun und Zentralafrikanische Republik. Von 1993 bis 1994 und im Jahr 1998 arbeitete er im DIPF (Deutsches Institut für Internationale Pädagogische Forschung) im Bereich Information und Dokumentation am Aufbau einer Datenbank zu Institutionen und Strukturen des Bildungswesens. Sein Interesse für angewandte Forschung im Sinne von Monitoring und Evaluierung entwickelte sich im Rahmen seiner Tätigkeit in Bildungsprojekten, speziell in Ruanda.