Универзитет у Источном Сарајеву
Филозофски факултет Пале

Посебна издања
НАУЧНИ СКУПОВИ

Књига 8
том 2/2

НАУКА И ГЛОБАЛИЗАЦИЈА

ФИЛОЗОФСКЕ НАУКЕ

ЗБОРНИК РАДОВА СА НАУЧНОГ СКУПА
(Пале, 17 - 19. мај 2013.)

Пале, 2014

НАУЧНИ СКУП

НАУКА И ГЛОБАЛИЗАЦИЈА

Књига 8
том 2/2

ФИЛОЗОФСКЕ НАУКЕ

Dejan M. Stanković[*]
Milja B. Vujačić
Institute for Educational Research
Belgrade

INTERNAL MONITORING AND EVALUATION OF IN-SERVICE TEACHER TRAINING PROGRAMS[**]

Introduction

In-service teacher training (INSET) is a potent way of improving teaching and learning. To fulfill this potential, training programs need to be of a high quality. Although quality assurance may vary across different educational systems, monitoring and evaluation are commonly seen as powerful tools for assuring quality in education. Monitoring can be defined as continuous data-gathering on ongoing activities or processes. Evaluation, on the other hand, is a judgment about the level of completion of certain standards, usually standards of merit and worth. Although evaluation might be understood as an informal reasoning, here we have in mind the kind of judgment that is more formal and systematic and based on some sort of research.

In the field of in-service teacher training, monitoring and evaluation should help us reveal how particular programs work, and with what results: whether we are accomplishing our goals, and how efficiently; what needs a complete change and what requires only a fine tuning, etc. Monitoring and evaluation programs can be implemented at different levels (from a central to a school level) and by actors with different functions within the INSET system: governing bodies; funding agencies; providers and participants (schools and teachers). With the recent trend of numerous services in education being delegated by the state to private agencies and civil society organizations, an old dilemma – how to arrange quality-assurance systems in education – has gained a new momentum. The local aspect of quality assurance grows in importance even more when central mechanisms are shown to be ineffective. Even if these mechanisms are sufficiently enhanced, their impact on quality improvement in INSET would remain limited against the potentials of local elements, especially in times when education quality becomes partly more decentralized, more contextual and variable.

All this leads to a growing awareness of the importance of the local role in defining and achieving quality.

In the field of INSET, this means that the genuine concern about program quality, as perceived by its providers, will lead to attempts of improvement. These attempts may involve internally initiated and organized monitoring and evaluation of training programs. We argue that these efforts are typically more effective and more reasonable when compared to remotely and centrally prescribed general measures. In fact, we would argue that schools and teachers should have an even bigger say in the quality assurance of INSET programs – after all their professional development is at stake here. The "advanced scenario" of an INSET quality assurance has it that teachers are supposed to make responsible and informed decisions as to whether participation in the program would be in the best interests of the school, teachers and students. An informed decision would then lead to a number of questions being posted to program providers, including asking what the effect of the program is so far, and what evidence is there to confirm this. So, in this "advanced scenario", a better market position would include those providers who are able to display the quality of their programs. In this, internal monitoring and evaluation performed by program providers may have a worthy marketing function, such as program promotion.

Internal monitoring and evaluation has other important functions, too. Data gathered in this way may guide efforts to improve a program in order to better accommodate the needs and goals of its participants (*program development*). Furthermore, those data can be used to report to licensing agencies and other stakeholders (*accountability*). Finally, monitoring and evaluation programs provide data that can be of *wider social importance*. Analysis of these data may point to possible problems in current educational practice, or indicate possible solutions which proved especially effective during the course of the INSET program. Those are the experiences and the knowledge that should be visible and present in policy and expert forums dealing with the development of education.

The local context

A new system of in-service teacher training in Serbia was introduced in 2003: staff development became obligatory; the offer of training programs became more liberalized; the choice of programs became more decentralized and autonomous; important elements of quality assurance were established such as an accreditation process, etc. The new system led to notable increase of attendance on teacher-training programs. Nevertheless, analysis suggests that the system faces serious problems, with a significant impact on its effectiveness (Stanković, 2011; Вујачић и др., 2011; Пешикан и др., 2010; Џиновић, 2009). It can be argued that the biggest problems of the current system are to be found in its quality assurance.

The Institute for Educational Research (in Belgrade, Serbia) is an organization which deals with educational research, and the dissemination and appli-

cation of its results as regards educational practice. It is one of the largest providers of in-service teacher training in Serbia. In 2011, the Institute offered seven nationally accredited training programs, duly following all mandatory guidelines for documenting and monitoring training programs. In some instances, it has made additional efforts to assess their merits. Nevertheless, there is no internal systematic approach to the documenting, monitoring and evaluation of its teacher-training programs.

The project

In 2011, the Education Support Program of the Open Society Institute commenced its international initiative 'Strengthening the evidence-based practice of educational civil-society organizations'. It has involved organizations interested in the joint development and field testing of data-gathering strategies for educational CSOs, in order to maximize 'the organizational and social value and utility of the data emerging from their daily work'. The initiative has offered the Institute for Educational Research an opportunity to take an in-depth look at its own undertakings, and possible ways to improve its monitoring and self-evaluation practices.

In the following sections, we will present the model of internal monitoring and evaluation of INSET programs that resulted from this project, its development and testing, and the lessons we have learned from these activities.

Development of the model

Efforts to create a new organizational approach for the monitoring and evaluation of teacher-training programs begun with the analysis of both national requirements for program evaluation, and the evaluative practices employed in that respect by the Institute for Educational Research. Both analyses sought to answer what kind of data were being gathering, in what manner, and for what purposes.

Although national guidelines imply using additional forms of evaluation, what is required from program providers is a one-page evaluation questionnaire and a written report from educators. The questionnaire measures teachers' perceptions and satisfaction with a given training program (including brief biographical data, 10 close-ended questions about the topics and methods, trainers, and working conditions, plus space for additional comments). The educators' report comments briefly on where the training was held, and answers to two open-ended questions about the positive aspects and possible difficulties of particular training. These requirements for program evaluation are very modest and they seem to be routinely executed. Their validity is not certain (Циновић, 2009) and is not clear that they have a real application (Пешикан и др., 2010).

The evaluative practices of the Institute were analyzed from a sample of two training programs. While both teams of teacher educators have followed na-

tional requirements, additional activities differ. One team used to have an additional evaluation questionnaire, and the team occasionally used training sessions for research purposes (Аврамовић и др., 2009; Аврамовић и Вујачић, 2010). The team of educators involved with the other training program had more comprehensive stance towards self-evaluation, devoting one training session to a sort of interactive evaluative discussion, using additional questionnaires, analyzing and publishing its results (Шефер и Радишић, 2010).

In parallel with these analyses, we have reviewed the literature in the field of program evaluation (Lancaster, 1983; Weiss, 1998; Frechtling, 2002) and more specifically, with regards to the evaluation of teacher-training programs (Guskey, 2000; Haslam, 2010). Furthermore, we have analyzed lists of guiding principles of several national evaluation associations (specifically Canada, the US, France, and Germany), particularly the program evaluation standards of the Joint Committee on Standards for Educational Evaluation (2011). The Committee's utility, feasibility, propriety, accuracy and accountability standards have been useful guidelines for thinking about varying issues of program evaluation.

The results of these analyses have provided a good foundation for outlining a draft model for the internal monitoring and evaluation of INSET programs.[1] This draft has been discussed within the organization, tested and continuously revised. The final model was published in a manual for program providers (Станковић и Вујачић, 2011).

The model

The model for the internal monitoring and evaluation of in-service teacher training programs is an organizing framework for varying types of activity, which providers might employ in order to review the quality of their training programs. Before describing the main parts of this framework, we will draw attention to its key features.

Key features

The model is principally intended *for program authors and providers*. For this reason, although the summative function is by no means neglected, *the formative purposes are central* to this model. In other words, the model puts *improvement before accountability*, even though the two are often inseparable. Further, the model assumes efforts which are primarily *internal* in relation to program providers. Although some of the incentives for using such a model might be of an external character, it is basically an internal endeavor – program providers appraising their own undertakings.

---

[1] Although, the initial project terminology favored the concept of data-gathering methodology, during the course of the project it became clear that more appropriate concepts would be those related to monitoring and evaluation.

Another core assumption of the model is that the complex interaction between a program and its *context* is the leading factor which affects a program's implementation and its results (see Figure 1). School conditions, policies and practices – as well as teachers' knowledge, attitudes, and motivation – forcefully shape a program's destiny, though high-quality programs do manage to accommodate and stimulate desired change.

Figure 1: Inter-relatedness of a school context and the training program
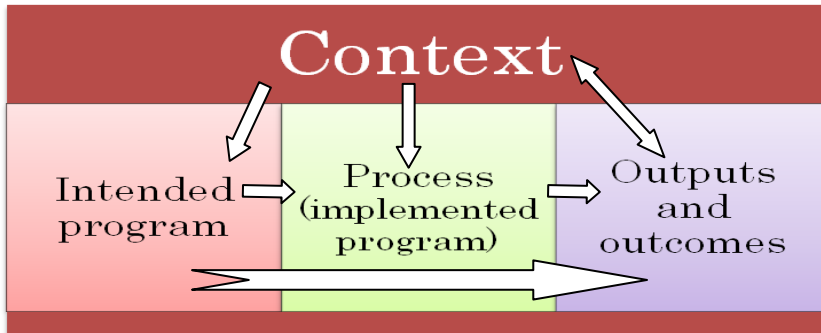
The model embraces a pragmatic stance in relation to data-gathering methodology. It favors *methodological pluralism*: suitable methods are those which provide valid and accurate data, regardless of individual methodological preferences. Hence, the model supports a mixed-method approach which rationally connects quantitative and qualitative methodologies. Generally speaking, all data about training programs come from three sources (see Figure 2). The first of these sources is people: teachers, students, parents, principals, etc. The second comprises products which, in a school context, may include organizational plans and reports, teacher and student portfolios, etc. The third data source is events: different pedagogical situations and interactions, such as classes, extracurricular activities, staff meetings, parent-teacher conferences, etc. Depending on the type of source, suitable data-gathering methods may include questionnaires and tests, interviews and focus groups, observations and content analysis.

Figure 2: Methodological framework
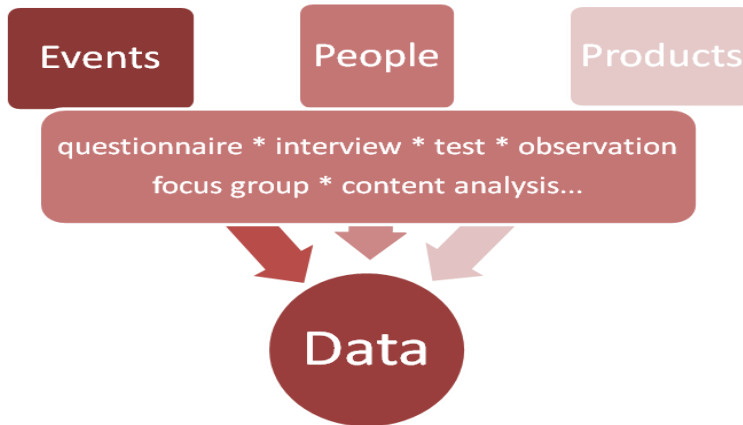
The remaining key feature of the model rests in the *integration of monitoring and evaluation* into a distinct instrument for organizational learning and development. In this way, monitoring – as a process of continuing data gathering – becomes the basic form of evaluative research. It does not, however, exclude the option of using additional forms of research designs: (quasi)experiments, *ex post facto* research, case studies, etc. However, there are a couple of reasons why this model puts forward monitoring as a main source of data about the program. Firstly, monitoring is highly appropriate for the evaluation process, since it includes data gathering which occurs during the program implementation (and even before it begins). This gives opportunities for immediate adjustments of the program to a particular group of participants. For instance, if gathered data show that participants would like to hear more real-life examples of presented innovations, then teacher educators could make use of this data for planning any further training sessions. These ongoing changes are not feasible with retrospective investigations, as is often the case, for example, with *ex post facto* research design. Monitoring program implementation, therefore, enables the basic function of evaluation: improving the program, even while it is still in progress.

The second reason for placing monitoring at the heart of evaluative research relates to the numerous and diverse contexts of program implementation, and to plentiful spontaneous and deliberate program variations. Research designs with a small number of groups of participants or schools under investigation, such as experiments and case studies, can not demonstrate how a program behaves in varied contexts, and how spontaneous or deliberate variations will affect program results. Monitoring, however, provides data on program implementation and program results from a wider diversity of contexts and realizations. This contributes to the validity and reliability of the data on which we base our evaluations.
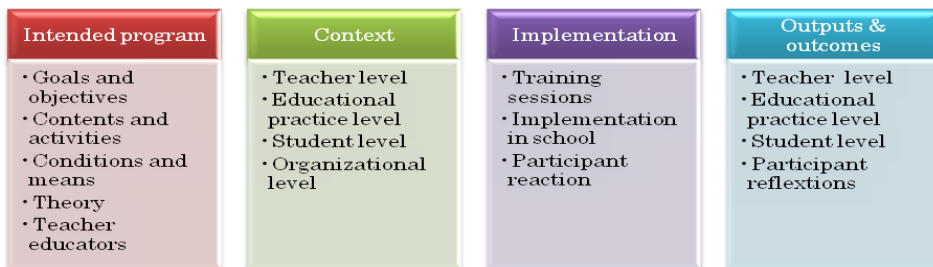
Internal monitoring may prove to be more cost-effective than other types of evaluative research, since it may be performed parallel to program implemen-

tation. Other research designs can be demanding as they require hiring external researcher or place the full emphasis of teacher training on a given research. This will de-focus program implementation, or even put it on hold temporarily – a luxury that many providers can not afford. On the other hand, even those program providers who are not skilled in research may with initial external help soon become sufficiently competent that monitoring ceases to challenge them.

A framework for monitoring and evaluation

Traditionally, the main questions that need to be addressed by evaluators are those related to program results. What are the program outputs and what are the outcomes? Have the program goals been achieved? What are the short-term and long-term effects of the program? Besides these questions, evaluators are increasingly expected to perform process evaluation as well, addressing such questions as the following. How does the program actually work? What activities have been carried out? What kinds of problems occurred? The framework model for the internal monitoring and evaluation of INSET programs naturally includes these paramount aspects (i.e. program implementation and its results). Because the model is intended for program authors and providers, thus having primarily self-evaluative and self-improvement purposes, it also places equal importance on gathering data on the contextual aspects and changes that the intended program is experiencing. In sum, the framework for comprehensive monitoring and evaluation includes data on the main segments of training programs: the intended program; its context and implementation; and outputs and outcomes.

Figure 3: A framework for monitoring and evaluation

Intended program	Context	Implementation	Outputs & outcomes
·Goals and objectives ·Contents and activities ·Conditions and means ·Theory ·Teacher educators	·Teacher level ·Educational practice level ·Student level ·Organizational level	·Training sessions ·Implementation in school ·Participant reaction	·Teacher level ·Educational practice level ·Student level ·Participant reflextions

Intended program

Any given program must have its own 'written and filed' description, but this is by no means a one-shot activity, because programs are modified during their own 'course of life'. Programs are affected by both spontaneous and deliberate changes. Sometimes, educators realize that – after each successive training session – some program goals have proven to be unrealistic; that some activities are redundant or ineffective; that some new training materials are especi-

ally useful. Presumably, these insights will lead to modifications which should be documented and carefully explained, so the basic documentation accurately represents the current shape of the intended program. This is important, since the intended program provides the basis for assessing what will come to be implemented (e.g. how far the implementation deviates from the initial plan) and for interpreting final outputs and outcomes. If the latter are not satisfactory, we should be able to discern whether the reasons are attributable to the plan or to its realization.

A detailed description of the program also serves to:

1) Identify the main elements that should be monitored and evaluated;

2) Explicate the program's quality standards; and

3) Explicate the links between the main elements of the program (overall program logic).

To illustrate this, we may imagine a program documentation that specifies an optimal number (e.g. 20) and the acceptable number of participants (10–30). Having more than 30 participants, according to the explicit assumptions of authors, would prevent interactive and experiential learning, while fewer than 10 participants would not form a 'critical' number needed to successfully implement the innovations within the school. Here, we may conclude that the number of participants is an important element that needs to be documented; quality standards of this element are made clear (the optimal and acceptable number), and the links between number of participants and opportunities for learning and change are explicated as well. In this way, a program description will direct the choice of elements to be monitored and evaluated; it will help us understand why are we monitoring and evaluating these elements; it will also teach us how we should interpret gathered data.

A description of the intended program should explicate: basic information (target groups, duration, etc); goals and objectives; program activities, with detailed explications of its content, duration, schedule, etc.; materials and resources needed for effective program implementation; the theoretical background; and the professional characteristics of teacher trainers.

Context

Gathering data about the school context has at least three important functions. Firstly, it makes it possible to adjust program goals and activities to the context of its implementation, as well as to the needs and goals of program participants. For example, data about pressing problems in relation to student behavior in a given school may help trainers tailor classroom management program to fit a particular school or teachers. Data on the context may indicate potential problems with program implementation, but also suggest possible ways for preventing and/or overcoming them.

Secondly, data about the local context may serve to determine the initial state which can be compared to 'final' state, after the program completion. Re-

sults of this comparison are some of the most important indicators of program effectiveness, both at the individual school level and the aggregate level. Ideally, this would allow the true effects of the program to be determined (i.e. the value that it adds). Analysis of the context data also provides an account of the type of school where the program has been implemented (where it had greater or lesser success), thus allowing a hypothesis of the potential effects of contextual factors on program effectiveness.

Finally, gathering data about the context allows varying educational issues for academic and policy purposes to be analyzed. For example, teachers' perceptions of their own cooperation with parents (e.g. gathered for a program which aims to develop an effective partnership between homes and school) are useful material for expert analysis and policy considerations in the field.

Contextual data will comprise data on teachers, students, educational practices and the overall school context. Nevertheless, the choice of aspects to be assessed should be selected according to the program (its goals, requirements, the target group, etc.).

Data on teachers may be related to their knowledge, skills, abilities, attitudes, beliefs, prior experience, competencies, goals, aspirations, motivation and expectations. Are the ideas conveyed by the program unfamiliar to teachers, or are they already well-known? What are teachers' experiences and competencies in the field? How motivated are teachers to take part in the program and what are their expectations?

Similarly, we would like to know about students' knowledge, skills, attitudes, competencies, behavior, etc. However, the choice should be guided by the features of the program. Most INSET programs have the ultimate goal of improving students' learning, knowledge, behavior or well-being. If one wants to establish a program's effects, the best way would be to gather data on these aspects before and after the program.

Depending on the program, the educational practice of interest might be at the level of the class or students (e.g. didactics, classroom management), at the school level (e.g. teacher collaboration, development planning), and beyond school practices (e.g. collaboration with parents and the local community). Measuring educational practices is a very demanding task, but it is especially necessary in those programs where the modification of educational practices is set as a chief indicator of a program's success.

The school context powerfully affects program implementation, so a detailed snapshot of the situation is of great importance. This set of data may comprise such aspects of the school as: size; resources; the immediate environment; policies; organization and management; climate; home–school collaboration, etc.

Implementation

Monitoring program implementation reveals similarities and differences between the intended and the implemented program; it helps us understand pro-

gram outputs and outcomes, and suggests ways of improving the program later on.

Teacher-training programs usually involve training sessions (seminars). They may be implemented in accordance to the initial plan to a greater or lesser extent. This is something that only program providers can be aware of. Therefore, they are the prime source of this type of data. Here, monitoring may require gathering data on: content (Were all intended topics covered? Were there any new topics?); activities (Was the number / schedule / duration of activities adequate? Were the participants active?); and conditions (Was the room spacious enough? Was there adequate equipment?).

On the other hand, program providers should be interested to learn how participants perceive varying aspects of training sessions, including: content (Was it relevant / useful / coherent / understandable / interesting / innovative?); working methods (Was it interactive / experiential / inspiring / relevant?); trainers (Were they clear / prepared / competent / communicative?); time management (Was the duration / structure / tempo satisfactory?); and conditions (Were the facilities / technical equipment / technical support / refreshments / working atmosphere satisfactory?).

Many training programs include a phase when teachers actively experiment and test new ideas in the real school context. Here, monitoring should pertain to data as to what actually happened and how. Moreover, it should include data on the organizational support. This may come from school management and colleagues, and may be of a material and non-material nature (including help, encouragement, and patience).

Outputs and outcomes

Monitoring and evaluating outputs and outcomes are complex endeavors. Nevertheless, they are so essential that extra effort is needed to gain at least basic data. In fact, what we are interested in most is whether our program gives good results, whether teachers and students are benefiting from it, whether it justifies the resources and trust. Evaluating outputs and outcomes also has an important formative function: it may indicate how to further improve the program.

Outputs, as immediate program results, are to be seen as more than easily documented figures on training sessions held, number of participants, etc. It should also be understood in terms of changes in teachers' competencies, knowledge, skills, attitudes and beliefs. Program outcomes are to be understood as its indirect results, consequences that can be attributed to program outputs. Outcomes could be identified with changes in educational practice and, more importantly, changes at the level of student knowledge and behavior induced by the program.

Outputs and outcomes can not usually be measured directly – they have to be established through the analysis and interpretation of relevant data. The best methodology would be to compare pre-program and post-program data whi-

le controlling for other important factors. It should be noted that a program may also lead to unintended results, which may be more or less desirable. That is why it is of great importance to be open to the varying signals that may arise from data.

Finally, we should record participants' and other beneficiaries' reflections on the program. Although these are not program results in a strict sense, they are included here because they should be gathered well after the program is completed. We are interested, for example, in how innovations have been sustained or altered, what benefits they brought and to whom, how the program looks from a distance, and how it could be improved.

Testing the model

The model for the internal monitoring and evaluation of in-service teacher-training programs was tested using a case-study design. The model was tested in one primary school on one of the Institute's training programs: 'How to think differently and in an interdisciplinary way'. This three-day school-based training program aimed to enhance teachers' competencies needed for thematic/interdisciplinary instruction and the development of children's creativity and lateral thinking (Шефер, 2005).

The whole set of data-gathering techniques was tailored to fit this particular training program. Just as the model assumed, each and every program provider should utilize their program in that the manner which best serves its intentions. Nevertheless, many of the questions used in questionnaires, focus groups and interview guides are not program specific, thus being potentially useful for monitoring and evaluating other teacher-training programs, too.

Testing began with focus groups of teachers and students; interviews were held with school principal and school advisers. This was done before the program implementation started.

Figure 4: Sample of questions for focus groups and interviews (prior to training)

For teachers:
What would you like to change in your teaching practice? Why?
Have you ever practiced thematic/interdisciplinary teaching? When? How?
What do you think of it?
What do you expect from the training? What do you want to learn from it?

For students:
What does the typical lesson look like? Do you like it?
Do you have opportunities to learn through games and research projects?
Do you learn together with your peers? Would you like to do that more often?

For school principals and advisers:
What do teachers in your school need to improve their teaching?
Are you familiar with the idea of thematic/interdisciplinary teaching?
Are you interested and motivated in its implementation at your school?
What do you expect from this training session?

At the end of the first training day, participants filled out the 'context' questionnaires. There were two versions, one for teachers and one for school principals. For both questionnaires, many of the questions were taken from large-scale international surveys (e.g. TIMSS, TALIS, PISA).

Figure 5: Indicators in the context-related questionnaires

Teacher questionnaire:
biographical data, job satisfaction, school climate, working conditions, student support, implicit pedagogy, professional collaboration, teaching practice, professional development needs, teacher self-efficacy

School principal questionnaire:
school size, school socio-economical environment, school facilities, school resources, school climate, student discipline, biographical data, management/leadership practices, implicit pedagogies

At the end of the second training day, a group discussion with all participants was held. The aim was to elicit their opinions on the training so far, what was difficult for them, and how this should be overcome.

Figure 6: Sample of questions for group discussion (during the training)

> *Were the training topics relevant for your practice?*
> *Were the working methods appropriate?*
> *Was your prior knowledge and experience appreciated?*
> *Were the educators well prepared? Clear? Open to questions?*

Between the second and the third training days, teachers had one month to carry out the thematic/interdisciplinary instruction they had jointly planned at previous training sessions. During this process, a couple of classes were visited, but without using any observation protocols. After the 'thematic week' was over, two interviews with individual teachers and a focus group with students were held in order to examine their opinions on the program implementation and its results.

Figure 7: Sample of questions for teachers and students (after program completion)

> **For teachers:**
> *How did you prepare thematic/interdisciplinary instruction?*
> *Did you have any new roles? What were they?*
> *What did your communication with students comprise?*
> *Did you learn anything new?*
> *Will you use this approach in future?*
>
> **For students:**
> *Could you tell us what you did in the classes as regards thematic instruction?*
> *Was that interesting for you?*
> *What did teachers do in these classes which differed from usual practice?*
> *Would you like classes to more often be organized in this way?*

Finally, this training program (as a rule) devotes the whole third day to presentations of what happened during the implementation of thematic/interdisciplinary instruction in the school. This gives opportunities for mutual feedback, reflection and in-depth discussion. At the end of the day, participants fill out the final evaluation questionnaire (mainly comprised of open-ended questions).

Figure 8: Sample of questions from the evaluation questionnaire

> *What was most difficult for students? What problems were there?*
> *What was most interesting for your students?*
> *Where were you personally most successful?*
> *Did you collaborate with your colleagues? To what extent?*
> *What has changed the most in your thinking about teaching after this training?*
> *Are students satisified with this type of instruction?*
> *Are students more succesful after having this type of instruction?*

Lessons learned

The model for the internal monitoring and evaluation of in-service teacher-training programs significantly improves the scope and the quality of data on the program, thus providing a solid foundation for drawing summative conclusions on the merit of the program overall (in all schools/contexts) and on each particular training session (in one school unit).[2] The formative function of the model is also very much emphasized: it is of great informative value for internal improvement efforts,[3] and potentially for the wider community (e.g. educational authorities).

The model is more ambitious then merely documenting training facts and measuring levels of participant satisfaction (the 'happiness quotient'). In fact, a full application of the model is feasible mostly with long-term INSET programs, probably those which are part of a larger organizational or system--improvement projects. Shorter programs would entail a more unassuming data-gathering approach. For example, if we want to assess how the teaching practice has changed, then instead of observing classes, we might ask students, teachers and other school professionals what they think about it.

Data gathering faces serious obstacles – two of which appear to have the greatest impact. Firstly, it seems that time is a scarce resource in schools, and a crucial resource for any data-gathering activity. Program participants might be very reluctant to set aside sufficient time. Secondly, although we will not argue that there is a general lack of awareness of the need to pay more attention to evaluation, the fact is that there is an underdeveloped culture of evaluation in Serbian schools.

---

[2] Teachers from the school used as a case study felt that the program was very good overall: useful, interesting, practical, innovative, increasing the motivation and creativity of students, etc.

[3] The case study clearly defined room for improvement: the theoretical part should be made more interesting if not shorter; educators should present more examples from other schools; educators should demonstrate new ideas by model teaching, etc.

Testing the model has provided a few messages to how these barriers might be overcome. The program participants should be informed thoroughly and in a timely manner about what is going to be monitored and evaluated, why and how, and for what the data will be used. Even more instrumental for mutual cooperation would be to involve program participants from the onset in designing a plan for monitoring and evaluation. This will offer the possibility for teachers to take a more active stance to data gathering as well (e.g. organizing data-gathering sessions themselves, recording the conclusions, and keeping other teacher trainers informed). This is also a way of appreciating the teachers' right to decide on important aspects of their professional development, including its evaluation.

Another lesson is that monitoring and evaluation should be efficient in relation to the resources needed (human, financial, and time resources). Therefore, it would be wise to think of different solutions to those activities which are usually time-consuming and expensive (e.g. instead of extra travel to school, teacher trainers should make the best use of new modes of communication, especially those that the Internet offers). Also, data-gathering activities should not be too demanding, redundant or tiresome. Educators should always take care not to jeopardize a program's implementation by its monitoring activities. They should avoid generating confusion or a tense atmosphere among participants which might be caused by constantly jumping from the role of trainer to that of evaluator. This leads to the next lesson learned: monitoring and evaluating needs to be carefully planned. But planning and carrying out data gathering and data analysis requires at least basic research competencies. In some instances, this will necessitate additional investments in building such capacities.

Internal monitoring and evaluation programs are not easy tasks. This is true for any authentic self-evaluative effort. Nevertheless, the potential benefits will outweigh the costs. This case study could not provide proof of this, but it definitely provided evidence that the model is a very promising tool for improving INSET programs. Now being disseminated in Serbia in the form of a manual for program providers (Станковић & Вујачић, 2011), we hope it will instigate new thinking on the quality of INSET programs in Serbia and the means for its improvement.

References

Аврамовић и др 2009: З. Аврамовић, Д. Станковић, и М. Вујачић, Квалитет дидактичких знања наставника. у: Ђ. Комленовић, Д. Малинић и С. Гашић Павишић (ур.): *Квалитет и ефикасност наставе,* Београд: Институт за педагошка истраживања, 127–140.

Аврамовић и Вујачић 2010: З. Аврамовић и М. Вујачић, *Наставник између теорије и наставне праксе*, Београд: Институт за педагошка истраживања.

Циновић 2009: В. Циновић, Изазови евалуације семинара на основу иску-
става наставника. у: Ђ. Комленовић, Д. Малинић и С. Гашић Пав-
ишић (ур.): *Квалитет и ефикасност наставе*, Београд: Институт за
педагошка истраживања, 321-333.

Frechtling 2002: J. Frechtling, *The 2002 User-Friendly Handbook for Project
Evaluation*, Arlington, VA: The National Science Foundation.

Guskay 2000: T. R. Guskay, *Evaluating Professional Development*, Thousand
Oaks, CA: Corwin Press.

Haslam 2010: B. M. Haslam, *Teacher Professional Development Evaluation
Guide,* Oxford, OH: National Staff Development Council.

Joint Committee on Standards for Educational Evaluation 2011: *The program
evaluation standards* (3rd edition.), Newbury Park, CA: Sage.

Lancaster 1983: F. W. Lancaster, *Guidelines for the Evaluation of Training C-
ourses, Workshops and Seminars*, Paris: UNESCO.

Пешикан и др 2010: А. Пешикан, С. Антић и С. Маринковић, Концепција с-
тручног усавршавања наставника у Србији – између прокламованог
и скривеног модела. *Настава и васпитање*, 59 (2), 278–296.

Шефер 2005: Ј. Шефер, *Креативне активности у тематској настави*, Б-
еоград: Институт за педагошка истраживања.

Шефер и Радишић 2010: Ј. Шефер и Ј. Радишић, Креативност и интердис-
циплинарност у нашој настави, *Настава и васпитање*, 59 (2), 205–
223.

Станковић 2011: D. Stanković, Sistem profesionalnog razvoja nastavnika u -
Srbiji: glavne teme i pravci razvoja. u: T. Vonta & S. Ševkušić (eds.): *Iz-
zivi in usmeritve profesionalnega razvoja učiteljev / Izazovi i usmerenja
profesionalnog razvoja učitelja,* Ljubljana: Pedagoški inštitut v sodelo-
vanju z Inštitutom za pedagoška istraživanja Beograd, 87-101.

Станковић и Вујачић 2011: Д. Станковић и М. Вујачић, *Праћење и вредно-
вање програма стручног усавршавања наставника: Приручник за
ауторе и реализаторе програма*, Београд: Институт за педагошка
истраживања.

Вујачић и др 2011: М. Вујачић, Ј. Павловић, Д. Станковић, В Циновић, и И.
Ђерић, *Представе о образовним променама у Србији*, Београд: Инс-
титут за педагошка истраживања.

Weiss 1998: C. H. Weiss, *Evaluation methods for studying programs and p-
olicies. (2ⁿᵈ edition),* Upper Saddle River, NJ: Prentice Hall.