

The Effectiveness of a Suggested Training Course in Developing EFL Teachers' Evaluation Practices at the Secondary Stage

Dr. Shery Eid Mohammed AbdulAal

Ph.D. degree holder in TEFL from

Faculty of Education – Zagazig University

Abstract

The present study aimed at examining the effectiveness of a suggested training course in developing EFL teachers' evaluation practices at the secondary stage. The study adopted the experimental design, using one experimental group and one control group. Participants of the study were 30 in-service EFL teachers in the academic year 2016/2017. They were randomly assigned into two groups, experimental ($n=15$) and control ($n=15$). The experimental group received training through the suggested training course, while the control group continued to study the content in the traditional way. This study tried to crystallize the prime and secondary elements of training courses in general and EFL training courses in particular, and the needs of EFL in-service teachers concerning evaluation practices. Therefore, the researcher designed a suggested course to teach evaluation practices. Results indicated that there was a statistically significant difference between the mean scores of the experimental group and those of the control group favoring the post scores. It was concluded that the suggested training course was effective in enhancing EFL in-service teachers' evaluation practices. Results of these findings suggest important implications for individuals who are required to review, design, or modify teacher evaluation programs.

Keywords: *In-service training course – Evaluation practices.*

Introduction

Evaluation in the context of education is a process used to obtain information from testing, from direct observations of behavior, from essays and from other devices to assess a student's overall progress towards some predetermined goals or objectives. It includes both a qualitative and a quantitative description and involves a value judgment of overall student behavior.

Evaluation and measurement are not the same, although evaluation involves measurement. If we assess a student's knowledge and understanding in a subject by means of an objective or an essay type test, that is measurement. If the teacher puts a value on the student's work, talents, attitudes and other characteristics of behavior that is evaluation. Evaluation should in part involve testing that is non-subjective on the part of the teacher; otherwise it is likely to be erratic and unreliable.

In the summer of 2010, a state Department of Education commissioned the Center for Innovative Technology (CIT) to support its teacher evaluation/performance models development project. During the project, utilizing the proven proprietary CIT Connect process, the CIT Connect project team compiled and synthesized current research, conducted data collection surveys to establish the current status of teacher evaluation within the state, and conducted interviews with representatives of departments of education in a number of states and localities to identify nationwide trends and models and to develop best practices recommendations. The CIT Connect team analyzed the teacher evaluation programs and practices at the state and district level in a number of states. The team found that the process of evaluating teachers is similar in most jurisdictions. Very few jurisdictions have mature programs (in place five years or more). Several have programs underdevelopment, a few are conducting pilot programs and a number plan to implement an evaluation system by the 2012-2013 school year.

To some extent trying to achieve improvement through accountability causes tensions. An emphasis on accountability may in some instances lead teachers to feel insecure or fearful and reduce their appreciation of their work (OECD, 2009). By contrast, teachers and their unions expect opportunities of social recognition of their work and opportunities for professional growth through the development of a formative system of teacher evaluation (Avalos and Assael, 2006).

Teacher evaluation for improvement purposes is likely to benefit from conditions such as:

- A non-threatening evaluation context.
- A culture of mutually providing and receiving feedback.
- Clear individual and collective objectives with regard to improving teaching within the school as well as a sharing of school objectives.
- Simple evaluation instruments such as self-evaluation forms, classroom observation, and structured interviews.
- A supportive school leadership.
- Opportunities to enhance competencies as well as resources and means to improve practice.
- Teacher evaluation integrated in a system of school self-evaluation and quality assurance.

In turn, teacher evaluation for accountability is likely to benefit from conditions such as:

- An independent and objective assessment of the teacher's performance.
- National-level standards and criteria across schools.
- An evaluation component external to the school and more formal processes.
- Well-established rules regarding the consequences of the evaluation.
- Clear individual objectives with regard to all aspects of a teacher's performance.
- Well-trained, competent evaluators of teaching performance.
- Impact on professional development plan.
- Possibilities for appeal for teachers who feel they have not been treated fairly.

Responsibilities for teacher evaluation

Responsibilities for teacher evaluation are typically shared between educational authorities, including quality assurance

agencies such as inspectorates, schools and their leadership, and teachers themselves.

The role of educational authorities

Educational authorities play a major role in the conception and application of teacher evaluation, since they set the national learning outcome objectives, agree standards for the teaching profession and establish the norms that regulate teacher evaluation. In some countries, they play a direct role in the implementation and monitoring of teacher evaluation procedures. This might include the design of specific evaluation tools and instruments, the determination of evaluation criteria, the distribution of evaluation duties, and the follow-up on evaluation results. In other countries, educational authorities establish general principles and guidelines only and give schools considerable leeway to adapt the teacher evaluation model to their particular circumstances (for instance, letting schools define evaluation criteria).

The role of inspectorates

In many countries education inspectorates take full responsibility for teacher evaluation. This includes developing teacher evaluation procedures and undertaking individual teacher evaluations with school inspectors taking the role of evaluators. In others, the inspectorate does not take responsibility for individual teacher evaluations but, instead, has an important role in stimulating both the quality of school leadership and the quality of teaching. This is typically done through feeding back the results of external school evaluation, mostly consisting of feedback on leadership and management, feedback on the quality of the teaching and learning processes, and feedback on school climate. In general, the Inspectorate also has an eminent role in modeling and disseminating good practice in teacher evaluation.

The role of schools and school leadership

The effective operation of teacher evaluation depends to a great extent on the way the concept and practice of school

leadership is established in schools. The role of school leaders in teacher evaluation differs across countries. In some countries, it may consist of the simple implementation of centrally-dictated regulations. In this case, real and recognized pedagogical leadership is necessary to use the teacher evaluation process developmentally and avoid the image of a bureaucratic device.

In other countries, school leaders take full responsibility for the evaluation of individual teachers. In Finland, the school director is the pedagogical leader, responsible for the teachers in her school and for the implementation of measures needed to enhance teaching quality. As a result, most of Finnish schools have a system that includes annual discussions aimed at evaluating the teachers' fulfillment of individual objectives set up during the previous year and determining developmental needs for the following year (UNESCO, 2007).

The complex interrelationship of personnel and career development and school improvement requires that the evaluation, management, and improvement processes are embedded and integrated in a system of school quality assurance.

The role of teachers

In some systems, teacher evaluation is based on reviews by peers, often more experienced and with more responsibilities. This is more typical of evaluation for improvement purposes. Evidence shows the power of developmental peer observation of lessons with clear foci and effective feedback in making a strong contribution to notching up the quality of instruction. When peers are the evaluators in accountability-driven teacher evaluation procedures, issues of legitimacy are particularly relevant to address.

Context of the problem:

Based on the results of the questionnaire in April, 2016 as a pilot study to identify the difficulties facing EFL teachers at Zagazig language secondary school (n=30), it was found out that EFL teachers lacked some evaluation practices skills. So, the

researcher has realized that in-service teachers of English are in need of training on the new techniques of evaluation practices instead of the traditional methods which have counted on the theoretical approach rather than the practical one.

Statement of the problem

In light of the aforementioned argumentative account, it is clear that the secondary school EFL teachers lack the skills of evaluation practices as has been revealed through the pilot study. It could also be mentioned that giving the activities of a suggested training course would help in achieving high standards of skill mastery for the EFL teachers. Accordingly, the problem under investigation in this study could thus be stated in the following main question:

How effective is a suggested training course in developing EFL in-service secondary school teachers' evaluation practices?

Questions of the study:

1. What are the evaluation practices skills needed for the EFL teachers at the language secondary school?
2. To what extent do those teachers master these skills?
3. How can a suggested training course be designed to enhance EFL teachers' evaluation practices?
4. To what extent is the suggested training course effective in developing EFL teachers' evaluation practices?

Significance of the study

The results of this study are hopefully expected to be useful to:

1. EFL teachers: This study may supply teachers of this stage with evaluation practices activities that could improve their performance in the EFL classrooms.
2. EFL students: This study may improve the EFL students' performance in practice and final exams.

Hypotheses of the study

1. There is no statistical significant difference between the experimental and control group in their performance on

the observation checklist pre conducting the training course.

2. There are statistically significant difference between teaching performance of evaluation practices of the control group and the experimental group in post observation, favoring the experimental group.
3. The suggested course is effective in developing evaluation practices of in-service EFL teachers at the secondary school.

Delimitations of the study

1. Thirty in-service EFL teachers at the secondary stage, in the In-Service Training Centre in Zagazig, Sharkia Governorate.
2. Evaluation practices skills which are specified by the questionnaire results.

Instruments of the study

The following instruments were designed and implemented by the researcher in order to achieve the research objectives :

- a. A questionnaire to identify the opinions of experts in the field of TEFL about the importance of evaluation practices to in-service EFL secondary school teachers.
- b. Performance observation checklist to pre and post observe the evaluation practice skills of EFL in-service teachers.

Review of literature

To be effective, a teacher assessment and evaluation system must employ a carefully constructed set of multiple measures. At NEA's Multiple Measures of Student Learning Working Group meeting in August 2010, Goe identified several specific benefits of such an approach. For example, she said the use of multiple measures strengthens teacher evaluation by providing a more complete picture of a teacher's contribution to student learning by increasing confidence in the results of a teacher's evaluation and by increasing the information available about collaboration for student success. Goe also said the use of a carefully constructed set of multiple measures contributes to teachers'

professional growth by creating opportunities for teachers to learn from their colleagues, and by providing teachers with greater insight into how their instruction is affecting student learning. Furthermore, the use of appropriate multiple measures sets the stage for improved teaching and learning by offering more complete evidence about student learning growth, particularly in areas where students are not tested by offering more complete evidence of learning growth for English language learners and students with disabilities, and by contributing to a comprehensive view of students' strengths and weaknesses.

Gathering multiple sources of evidence about teacher practice meets the need for accuracy and fairness of the evaluation process, taking into account the complexity of what a good teacher should know and be able to do. A range of instruments and information sources are typically used to evaluate teachers.

Teaching practices and evidence of student learning are likely to be the most relevant sources of information about professional performance. As a result, teacher evaluation is typically firmly rooted in classroom observation. Most key aspects of teaching are displayed while teachers interact with their students in the classroom. Classroom observations are the most common source of evidence used in OECD countries, whether American (e.g. Canada, Chile, United States), European (e.g. Denmark, France, Ireland, Spain) or Asian-Pacific (e.g. Australia, Japan, Korea). Such instrument shows whether the teacher adopts adequate practices in his more usual workplace: the classroom (Isoré, 2009; UNESCO, 2007).

Most teacher evaluation models require the individual teacher to set performance objectives for a given period of time in agreement with the school management. The evaluation then assesses the extent to which such objectives were met. The setting of objectives, as well as the evaluation itself, typically involves individual interviews which are an opportunity to trigger critical reflection between evaluators and teachers.

Another common instrument used in teacher evaluation is teacher self-evaluation. The perspective of the teacher being

evaluated is essential, because it allows teachers to express their own views about their performance, and reflect on the personal, organizational and institutional factors that had an impact on their teaching.

An instrument which typically complements teacher self-evaluation is a teacher portfolio, providing evidence about key aspects of the teacher's teaching. Different elements can compose teacher portfolios, including: lesson plans and teaching materials, samples of student work and commentaries on student assessment examples, teacher's self-reported questionnaires and reflection sheets (see Isoré, 2009, for a discussion on the use of portfolios including for improvement purposes versus accountability purposes). It should be noted that portfolios are not only a tool for evaluation, but also play a role in supporting a reflective approach to teaching practice that is a hallmark of effective teachers.

Comprehensive models of teacher evaluation involve the use of standardized forms to record teacher performance across the aspects being evaluated. These are key tools used in the evaluation process which seek to capture performance across the range of domains under evaluation.

a. Teacher testing

In some countries, for particular purposes such as the access to a permanent position or entry into the profession, teachers are the subject of testing to assess their general and subject-matter competencies. In some instances, the results of such tests can be used for teacher evaluation.

b. Student results

Student results are not commonly used as sources of evidence for teacher evaluation in countries (OECD, 2005; UNESCO, 2007). Given that a wide range of factors impact on student results, identifying the specific contribution of a given teacher is faced with numerous statistical challenges (Isoré, 2009). In this respect, the development of “value-added” models represent significant progress as they are designed to control for

the individual student's previous results, and therefore have the potential to identify the contribution an individual teacher made to a student's achievement. However, in order to be effective, value-added models require vast amounts of data to be collected through large scale national-level student testing across levels of education and subjects, an option with prohibitive costs.

The success of any teacher evaluation system greatly depends on the in-depth training of the evaluators. Experience suggests that evaluators should have a range of characteristics and competencies, including: (i) background in teaching, (ii) knowledge of educational evaluation theories and methodologies, (iii) knowledge of concepts of teaching quality, (iv) familiarity with systems and procedures of educational and school quality assurance, including the role of teaching quality in school quality and the role of teaching quality in personal development, (v) understanding of instrument development, including reliability and validity of observation and other assessment tools, (vi) awareness of the psychological aspects of evaluation, (vii) expertise with the quantitative rating of an assessment, and (viii) mastering the evaluation-related communication and feedback skills. These areas should receive priority for specialized training addressed to evaluators.

Given their crucial pedagogical role, school directors should benefit from dedicated training. Offers targeted at school leaders typically focus on human resources development and school quality assurance, including school self-evaluation. This usually involves personnel management, including aspects such as structured interactions with teachers, setting of objectives, linking school objectives to personnel development plans, making use of various sources of information on teaching quality and functional performance, development of instruments, and management instruments to use evaluation results.

Guaranteeing that teachers are provided with support to understand the evaluation procedures and to benefit from evaluation results is also vitally important. Teachers can benefit from training modules so they know what is expected from them

to be recognized as good teachers, and to be prepared to make the best use of the feedback received. This contributes for teachers to appropriate the process through support and coaching. Evaluation and feedback ought also to be important aspects offered in initial teacher education and regular professional development activities.

For the reasons previously mentioned, new methods in teacher's evaluation have appeared in order for some processes to take place.

1. Emphasizing the function of teacher evaluation to seek out, document, and acknowledge the good teaching that already exists.
2. Evaluating to reassure teachers and audiences (parents, legislators) that good work is going on.
3. Placing the teacher at the center of evaluation activity.
4. Using more than one person to judge teacher quality and performance.
5. Limiting administrator judgment role.
6. Using multiple data sources to inform judgments about teachers.
7. Using variable data sources to inform judgments about teachers.
8. Spending the extensive time and other resources needed to recognize good teaching.
9. Correctly using research on teacher evaluation; examine "research based" claims.
10. Attend to the sociology of teacher evaluation.
11. Use the results of teacher evaluation to encourage personal professional dossier building.

Method

The experimentation followed the steps below:

1. The researcher developed the study instruments and had them validated. Then, they were administered on the study participants before implementing the training course.
2. The researcher designed the proposed course, and had it validated by a number of TEFL experts. The course

consisted of 10 sessions. Each session is concerned specifically an aspect of evaluating students under three general headings:

- a. Teacher made tests.
- b. Other evaluation devices.
- c. Continuous evaluation or assessment.

Teacher made tests:

- *The essay test:* This type of test is helpful in measuring the students' ability to organize, interpret, evaluate and apply knowledge. It helps the teacher assess the students' ability to summarize, outline and see relationships and trends. It also tests the students' ability in fluency and clarity of expression.
- *The objective test:* The common characteristics of all types of objective tests is that there is only one correct answer to a question, unlike the essay type answer in which there are shades of right and wrong. There are many types of objectives tests, the four most commonly used objectives tests are: True/false, Completion, Multiple choice, and Matching
- *Oral tests:* Oral tests or quizzes in the classroom to assess students' learning. Most teachers use the questioning technique in their lessons and the quiz is simply an extension of this technique.

Other evaluation devices:

There are at least three other ways in which the evaluation process could be made more complete and these involve the use of *Assignments, Observation of in class and out of class behavior, and Individual discussions with students.*

1. Each session includes a group of activities. The number of activities is different from one session to another according to the objectives to be achieved. Some activities took 45 minutes while others took 60 minutes. Thus, each lesson took time according to the numbers of its activities including their presentation and procedures. Evaluation

of each lesson was at the end of implementing the whole activities of this session

2. 4. The researcher taught the experimental group by using the suggested course following training steps as follows:
 - a. *Presentation*: The researcher used various styles to initiate the practice of the activities whether through giving an introduction about the topic or asking questions or making discussions with the teachers.
 - b. *Practice*: At this stage, the researcher provided teachers with various activities through group discussion; teachers were given the opportunity to engage in practicing these activities with the instructor and among themselves.
 - c. *Evaluation*: At the end of each session, teachers were evaluated using questions about the chapter to measure the teachers' ability to master the evaluation practices.
3. 4. Having finished the training course, the researcher administered the study instruments. Data were collected and analyzed statistically.

Statistical Analyses and Results

To test the first hypothesis of the study which states that there is no statistical significant difference between the experimental and control group in their performance of the observation checklist before conducting the training course, Wilcoxon Ranks Test has been used to find out whether there are significant differences.

Table . Wilcoxon Signed Ranks Test for comparing the control group (N=15) and the experimental group (N=15) in the pretest of evaluation practices

Ranks	Mean Rank	Sum of Ranks	Z	Sig.
Negative	4.50	22.50	-0.707	0.480
Positive	4.50	13.50		

-Not significant.

As indicated in Table 1, Wilcoxon Signed Ranks Test revealed that the differences between the mean ranks of the control group and the experimental group on the pretest of teaching performance were not significant. ($Z=-0.707$, $p>0.05$). This means that both groups of the study, the control and the experimental, were equivalent before teaching the training course proposed.

To test the second hypothesis of the study which states that "there are statistically significant difference between teaching performance of evaluation practices of the control group and the experimental group in post observation, favoring the experimental group, "Wilcoxon Signed Ranks Test has been to find out whether there are significant differences.

Table 2. Wilcoxon Signed Ranks Test for comparing the control group (N=15) and the experimental group (N=15) in the posttest of evaluation practices

Ranks	Mean Rank	Sum of Ranks	Z	Sig.
Negative	8.00	120.00	-3.412	0.001*
Positive	0.00	0.00		

*Significant ($p<0.05$).

As indicated in Table 2, Wilcoxon Signed Ranks Test revealed that the differences between the mean ranks of the control group and experimental group on the posttest of teaching performance were significant. ($Z=-3.412$, $p<0.05$). This means that the gain scores of the participants of the experimental group were greater than those of the participants of the control group.

To test the third hypothesis of the study which stated that "the suggested course is effective in developing evaluation practices of in-service EFL teachers at the secondary school, the ratio of acquisition average of Blakes has been used.

$$\text{Ratio of acquisition average of Blakes} = \frac{A-B}{D-B} + \frac{A-B}{D}$$

A= Post average

B= Pre-average

D= Maximum score

Teachers' performance average in post- observation was 93.0000

Maximum score= 100

Pre-average= 54.3333

Post- average= 93.0000

So the ratio of acquisition average = $0.846 + 0.386 = 1.232 \sim 1.2$

The ratio of acquisition average = 1.2

These results indicate that the training course is very effective in teachers' performance according to Pachman, D. Clear, A., Mayer, T. (1971: 472-473)

Discussion and interpretation of the results

Making the best use of teacher evaluation results raises a number of challenges such as feeding information on performance back to those who deliver educational services, most notably teachers and school principals, developing mechanisms for the improvement of teaching practices, notably teacher professional development; establishing rewards, support systems and consequences that flow from evaluation results, and developing the channels which ensure that information generated by teacher evaluation is used for educational policy development. These challenges highlight the importance of knowledge management in any teacher evaluation framework.

The following two key aspects are analyzed in turn:

Formative use of results

A key objective of teacher evaluation is to identify areas of improvement for individual teachers, leading to the preparation of individual improvement plans (including professional development) which take into account the overall school development plan. Without a link to professional development

opportunities, the evaluation process is not sufficient to improve teacher performance, and as a result, often become a meaningless exercise that encounters mistrust – or at best apathy – on the part of teachers being evaluated (Danielson, 2001; Milanowski and Kimball, 2003; Margo et al. 2008). Typically, schools are provided with autonomy to determine how teacher evaluation results feed into teacher professional and school development plans. School pedagogical leadership plays the key role in ensuring the effectiveness of such link (Pont et al., 2008). Another key element are the resources made available for professional and school development.

Summative use of results

Evaluation of teacher performance can also be used to determine career advancement, award performance rewards or establish sanctions for underperforming teachers. It constitutes an opportunity to recognize and reward teaching competence and performance, which is essential to retain effective teachers in schools as well as to make teaching an attractive career choice (OECD, 2005). In general, teacher evaluation frameworks might have consequences at the following levels:

Career advancement

Most countries do not link directly teacher evaluation results with teacher pay but, instead, to career progression (therefore establishing an indirect link with salaries). An example is England where teachers who meet the standards for “Post Threshold, Excellent and Advanced Skills Teachers” also access the higher pay scale (TDA, 2007a). Most teacher evaluation models relate results to the speed at which the teacher progresses in the career. Typically, ratings in the top categories of the rating scale make the teacher progress faster in the career scale while ratings in the bottom categories of the rating scale lead to no career progression during the period associated with the evaluation.

- *Decisions at key points in career.* The results of teacher evaluation can be used to make tenure decisions at the end of the probationary period, contract renewal

decisions for contract decisions and might influence the chances of an application to a given school post.

- *Performance rewards.* Teacher evaluation results might be used to award rewards to teachers. Typical rewards include the award of a one-off monetary prize (bonus pay), time allowances and sabbatical periods, opportunities for school-based research, support for post-graduate study, or opportunities for in-service education. In some instances the focus of the rewards is on group recognition and rewards are at the school or grade level rather than individual level. The “bonus” pay element should be approached with considerable caution. The evidence of the overall impact of such extra payments is mixed and can be contentious and potentially divisive (OECD, 2005).
- *Sanctions for ineffective teachers.* Some countries provide for sanctions for ineffective teachers beyond the standard consequences for career progression. For example, if underperformance persists following a number of evaluations, sanctions might include the removal from teaching duties (for teachers with civil servant status, this might translate into other functions within the school or another career within the civil service), or simply the termination of the contract. But early identification of underperformance is typically accompanied by a plan for in-service training for the improvement of practice.

However, it needs to be kept in mind that the *issues surrounding developing a closer relationship between teacher performance and reward are controversial in all countries*; and research in this field is difficult and has produced mixed results. There seems to be agreement that the design and implementation of performance-based rewards are crucial to their success. As explained in Harvey-Beavis (2003), there is a wide consensus that previous attempts at introducing performance-based reward programmes have been poorly designed and implemented (Mohrman *et al.* 1996; Ramirez,

2001). Problems in developing fair and reliable indicators, and the training of evaluators to fairly apply these indicators have undermined attempts to implement programmes (Storey, 2000). One problem identified is poor goal clarity because of a large number of criteria, which restricts teachers' understanding of the programme and makes implementation difficult (Richardson, 1999). Explanations of how, and on what criteria, teachers are assessed may be difficult to articulate. When this occurs, it is almost impossible to give constructive feedback and maintain teacher support for the programme (Chamberlin *et al.*, 2002).

Conclusions

The results of the present study revealed that:

Promoting teacher evaluation is clearly in the national interest as well as serving students and their families and communities. Teachers need feedback on their performance to help them identify how to better shape and improve their teaching practice and, with the support of effective school leadership, to develop schools as professional learning communities. At the same time, teachers should be accountable for their performance and progress in their careers on the basis of demonstrated effective teaching practice.

Developing a comprehensive approach may be costly but is critical to conciliate the demands for educational quality, the enhancement of teaching practices through professional development, and the recognition of teacher knowledge, skills and competencies. The expectation is that teachers engaging in reflective practice, studying their own methods of instruction and assessment, and sharing their experience with their peers in schools, becomes regular a routine part of professional life.

Study implications

There are multiple challenges to successful implementation. Implementation difficulties may arise as a result of a wide range of factors. First, there might be a defensive culture of evaluation, little tradition of peer evaluation, feedback and sharing of teaching practices. This might lead to some

resistance to application of teacher evaluation by particular groups in the school system. Second, there might be some technical challenges to implementation, including limited professional expertise of those with responsibility to evaluate, insufficient preparation of schools to implement evaluation procedures, limited understanding by teachers of the purposes and uses of evaluation, a sense of unfairness by those teachers being evaluated, the excessive workload inherent to the evaluation process, and the reluctance of many teachers to accept the legitimacy of the evaluators. Third, there might be issues of lack of resources for aspects of the teacher evaluation procedures, particularly the time needed for developmental work, observational evaluation and feedback.

While comprehensive teacher evaluation models - *e.g.* with the multiplication of instruments and evaluators - are more likely to provide a solid basis to evaluate teachers, limited resources make trade-offs inevitable. As explained in Isoré (2009), comprehensive teacher evaluation procedures imply greater direct and indirect costs at every stage of the process: reaching agreements on the design of the system requires time for discussions and consultations with all stakeholders (Avalos and Assael, 2006), training evaluators is expensive and requires time (Danielson, 1996, 2007), conducting evaluation processes induces additional workload for both teachers and evaluators, unless offsetting is made by reducing workload with other responsibilities (Heneman *et al.*, 2006), and aligning broader school reforms such as professional development opportunities requires more educational resources (Heneman *et al.*, 2007; Margo *et al.*, 2008).

It is essential to analyze the challenges of implementation of teacher evaluation. This includes reconciling the diverging interests of stakeholders, carefully analyzing policy alternatives and their likely impact and discussing them with stakeholders to aim towards consensus.

Adverse effects are particularly prone to occur when consensus has not been reached on the objectives for teacher

evaluation, its importance for the performance of the school system, and the practical options for implementation. Consensus is all the more precious to reach since local actors such as teachers and school leaders may be in the best position to foresee unintended consequences or judge what is feasible in practice, and since the effectiveness of teacher evaluation heavily depends on their cooperation.

Since teacher evaluation is eminently a technical matter and has a lot to benefit from worldwide evidence on best policies and practices, some countries bring together educational researchers and distinguished teachers into an advisory group to monitor and guide the implementation of teacher evaluation. Such group is in a good position to recognize good evaluation practices, to keep abreast of relevant research developments and, as a result, to provide advice based on sound evidence.

Teacher evaluation and the resulting feedback, reflection and professional development will only work if teachers make it work. To a great extent, it is the motivated teacher who ensures the successful implementation of reforms in schools. Hence, it is imperative to find ways for teachers to identify with the goals and values of teacher evaluation arrangements and practices (OECD, 2006). Teachers must be supported in understanding what the evaluation expects from them to be recognized as good teachers and in preparing adequately for the evaluation process.

References

- Avalos, B. and J. Assael (2006). Moving from resistance to agreement: The case of the Chilean teacher performance evaluation. *International Journal of Educational Research*, 45 (4-5) 254-266.
- Castro Ramos, C. (2009). *Teacher Evaluation in Portugal: Country Background Report for OECD*, Ministry of Education, Lisbon (available from www.oecd.org/edu/teacherevaluationportugal).
- Chamberlin, R.; Wragg, T.; Haynes, G. and Wragg, C. (2002). Performance-related Pay and the Teaching Profession: a

- Review of the Literature. *Research Papers in Education*, 17(1), 31-49.
- Danielson, C. (1996, 2007), *Enhancing Professional Practice: a Framework for Teaching*, 1st and 2nd editions, Association for Supervision and Curriculum Development (ASCD), Alexandria, Virginia.
- Danielson, C. (2001). New Trends in Teacher Evaluation. *Educational Leadership*, 58 (5),12-15.
- Danielson, C. and McGreal, T. (2000). *Teacher Evaluation to Enhance Professional Practice*, Association for Supervision and Curriculum Development (ASCD), Alexandria, Virginia.
- Economic Policy Institute (2010). *Problems with the use of student test scores to evaluate teachers*. Retrieved August 2010, <http://www.epi.org/publications/entry/6276>.
- Goe, L. and Croft, A. (2009). *Research -To-Practice Brief: Methods of Evaluating Teacher Effectiveness*. Washington, D.C.: National Comprehensive Center for Teacher Quality.
- Harvey-Beavis, O. (2003). Performance-Based Rewards for Teachers: A Literature Review. Paper distributed at the third workshop of participating countries on OECD Activity *Attracting, Developing and Retaining Effective Teachers*, 4-5 June, Athens, Greece. Available from www.oecd.org/edu/teacherpolicy
- Heneman, H.; Milanowski, A.; Kimball, S.; Odden, A. (2006). Standards-Based Teacher Evaluation as a Foundation for Knowledge- and Skill-Based Pay. *Consortium for Policy Research in Education (CPRE) Policy Briefs RB-45*.
- Heneman, H., A. Milanowski and S. Kimball (2007). Teacher Performance Pay: Synthesis of Plans, Research, and Guidelines for Practice. *Consortium for Policy Research in Education (CPRE) Policy Briefs RB-46*.
- Isoré, M. (2009). Teacher Evaluation: Current Practices in OECD Countries and a Literature Review. *OECD Education*

Working Paper No.23, OECD, Paris. Available from www.oecd.org/edu/workingpapers

- Jacob, B. and L. Lefgren (2005). What Do Parents Value in Education: an Empirical Investigation of Parents. Revealed Preferences for Teachers, *NBER Working Paper n°11494*.
- Little, O; Goe, L.; Bell, C. (2009). *A Practical Guide to Evaluating Teacher Effectiveness*. Washington, D.C : National Comprehensive Center for Teacher Quality.
- Margo, J., M. Benton, K. Withers and S. Sodha (2008). *Those Who Can?*, Institute for Public Policy Research (IPPR) Publications.
- Milanowski, A.; Kimball, S. (2003). The Framework-Based Teacher Performance Assessment Systems in Cincinnati and Washoe. *CPRE Working Paper Series TC-03-07*.
- NEA Foundation for the Improvement of Education (2000). *Engaging Public Support for Teachers'Professional Development*. Washington, D.C.: National Education Association.
- Mohrman, A.; Mohrman, S.; Odden, A. (1996). Aligning Teacher Compensation with Systemic School Reform: Skill-Based Pay and Group-Based Performance Awards. *Educational Evaluation and Policy Analysis*, 18(1), pp. 51-71.
- OECD (2005). *Teachers Matter: Attracting, Developing and Retaining Effective Teachers*. OECD, Paris.
- OECD (2006). The Teaching Workforce: Meeting Aspirations and Enhancing Motivation. *Education Policy Analysis 2005*, OECD, Paris.
- OECD (2009a). *OECD Review on Evaluation and Assessment for Improving School Outcomes: Design and Implementation Plan for the Review*, OECD, Paris [OLIS Document EDU/EDPC(2009)3/REV1].
- OECD (2009b). *Creating Effective Teaching and Learning Environments: First Results from TALIS*, OECD, Paris.

- Pont, B.; Nusche, D.; Moorman, H. (2008). *Improving School Leadership, Volume 1: Policy and Practice*, OECD, Paris.
- Ramirez, A. (2001). How Merit Pay Undermines Education. *Educational Leadership*, 58(5), pp. 16-20.
- Richardson, R. (1999). *Performance Related Pay in Schools: An Assessment of the Green Papers, A report prepared for the National Union of Teachers*. The London School of Economics and Political Science.
- Rand Education Research Brief (2004). *The Promise and Peril of Using Value-Added Modeling to Measure Teacher Effectiveness*. Santa Monica, Ca.: Rand Education.
- Santiago, P.; Roseveare, D.; van Amelsvoort, G.; Manzi, J.; Matthews, P. (2009). *Teacher Evaluation in Portugal: OECD Review*, OECD, Paris (available from www.oecd.org/edu/teacherevaluationportugal).
- Storey, A (2000). A Leap of Faith? Performance Pay for Teachers. *Journal of Education Policy*, 15(5), pp. 509-23.
- Stronge, J. and Tucker, T. (2003). *Handbook on Teacher Evaluation: Assessing and Improving Performance*, Eye On Education Publications.
- TDA (Training and Development Agency for Schools, 2007a), "Professional Standards for Teachers: Why Sit Still in Your Career?", United Kingdom.
- TDA (Training and Development Agency for Schools) (2007b). Professional Standards for Teachers, The Training and Development Agency for Teachers: London, www.tda.gov.uk/upload/resources/pdf/s/standards_a4.pdf
- UNESCO (2007). *Evaluación del Desempeño y Carrera Profesional Docente: Una panorámica de América y Europa*, Oficina Regional de Educación para América Latina y el Caribe, UNESCO Santiago, 2007.

