EFL Teacher Evaluation: A Theoretical Perspective

Khadija Alamoudi and Salah Troudi

Abstract In this chapter, we theoretically tackle the issue of evaluating English language teachers in educational institutions. In some parts of the world and out of all structural elements in institutes of education, teachers seem to receive the least amount of support and/or opportunities to assess their proficiency or to evaluate their teaching skills. Where teachers' performance evaluation does receive sufficient attention, cases sometimes can be expected where educators might be unaware of the latest up-to-date information concerning the methods that are utilized or the purposes behind the evaluation systems in their context and any policies related to it. The objectives of this chapter are to review and contribute to the current debates on the purposes and methods of evaluating language teachers' performance.

Keywords Teacher evaluation · Evaluation purposes · Evaluation methods · Language teacher evaluation

1 Introduction

Teachers are known to have a very significant influence on their students' achievement and to raise the interest of the pupils in the subject they are teaching. It is for that reason that high quality teaching is the goal of all language teachers. This might lead us to the importance of the evaluation of teachers and specifically EFL teachers in our case. Danielson (2001) assures that educators have realized that a well-designed system of evaluation is needed in order to improve their educational practices and to ensure a standard quality of teaching. Therefore, this chapter will be

K. Alamoudi (✉)
English Language Institute, King Abdulaziz University, Jeddah, Saudi Arabia
e-mail: khaalamoudi@kau.edu.sa

K. Alamoudi · S. Troudi
Graduate School of Education, University of Exeter, Exeter, UK
e-mail: s.troudi@exeter.ac.uk

useful to understand the current situation of EFL teacher evaluation and to offer insights to improve the existing practice of the system of evaluation.

2 The Concept of Evaluation in Education

In the field of education, many attempts have been made to clarify the concept of evaluation and to distinguish between evaluation and other closely related concepts such as measurement or assessment. In the following section, we will provide some of the established definitions for the term "evaluation" and how it has emerged in the field of education in order to distinguish it from the other related concepts.

2.1 Evaluation: Operational Definitions

As an educational concept, evaluation has received much attention in the literature and many definitions have been provided in order to help people conceptualize this significant notion. Into the context of evaluation, Ralph W. Tyler, a leading figure in educational evaluation, associates evaluation with the concept of objectives. According to Tyler (1950), evaluation is "the process of determining to what extent educational objectives are actually being realized" (p. 69). His objectives model had a lasting impact on evaluation conceptions. However, the model was criticized for the inability to present a method to assess educational objectives themselves. Cronbach (1963, as cited in Verma & Malick, 1999), on the other hand, links evaluation to decision-making instead of objectives and defines it as "the collection and use of information to make decisions about an educational program" (p. 47). His work involves evaluation in three different layers of educational decisions: Administrative regulation, course improvement, and decisions about individuals.

Although Cronbach's definition seems effective in guiding decision making, his model was criticized for equating evaluation to only one of its various roles. Another definition is provided by Rossi, Lipsey and Ferma (2004) who identify the concept of evaluation as the "use of social research methods to systematically investigate the effectiveness of social intervention programs to improve social conditions" (p. 16). Having a systematic method of evaluation is vastly considered in their definition; however, it seems that the developmental-based approach would benefit the most out of their model of evaluation. The idea of the systematic tactic has been taken further by Patton (2008) who defines evaluation as the "systematic collection of information about the activities, characteristics, and results of programs to make judgments about the program, improve or further develop program effectiveness, inform decisions about future programming, and/or increase understanding" (p. 38). Patton has provided not only a systematic method in his definition but the definition also embodies an inclusive description for various purposes. While all the previously mentioned definitions differ in their details and the ways

they conceptualized the term "evaluation", the decision to choose one of the definitions may depend on some other important factors, such as the evaluation context, research questions, and the issues to be addressed.

This chapter is informed by the last definition offered by Patton (2008) for a number of reasons. First, the definition is comprehensive in the sense that it includes a variety of purposes. Second, Patton considers evaluation as a systematic way to collect information about different aspects.

2.2   Emergence of Evaluation in Education

In the field of education, it seems that there is a consensus that the history of evaluation began before the turn of the 20th Century (Glasman, 1986; Guba & Lincoln, 1981; Norris, 1990). Glasman argues that the history of educational evaluation can be divided into three distinct phases: The first continued until the 1930s, the second lasted until the 1960s and the third is still going on. It seems that expansion rather than substitution of the old ideas is the main characteristic of the development of educational evaluation throughout those three periods. Evaluation was seen first as measurement in education and the focus was initially on the level of intelligence measurement for learners and their ability to learn a specific subject (Glasman, 1986). Glasman claims that educational evaluation before the 1930s was used widely in the life and physical sciences. On the other hand, Guba and Lincoln (1981) argue that during the last decade of the ninetieth century, Joseph Rice who is known as the father of educational research devised some achievement tests supporting his debate about the insufficient use of school time. His published test in 1904 has become the base for almost all tests that measure intelligence since then. However, the publication of Fredrick Taylor's *The principles of scientific management*, can be considered as the core effect of the ideas about standardization and systematization on industry which offers a systematic methodology for educational administration (Norris, 1990). Despite the fact that Ralph Tyler's contribution in the field of educational evaluation in the 1930s keeps evaluation synonymized with measurement, he is regarded by many as the father of educational evaluation and the invention of the term "evaluation" was attributed to him (Norris, 1990). This idea was opposed by Guba and Lincoln (1981) who argue that Tyler's method of evaluation has a distinctive advantage over the measurement-directed methods that were popular at that time. The reasoning in Tyler's approach is systematic in nature. This can be true given that Tyler's focus was on refining of programs and curricula in particular by means of examining educational objectives that can be considered as an essential impetus for evaluation.

2.3   The Changing Landscape of Teacher Evaluation

Medley, Coker and Soar (1984) briefly depict the teacher evaluation change of the twentieth century. They divide it into three main phases: (1) Questing for Great Teachers; (2) Determining the Quality of Teachers by Students' Learning; and (3) Observing Teaching Performance. In 1896, the issue of Great Teachers was evoked with a study conducted by Kratz who asked 2411 students in Iowa to define the features of the best teachers (Medley, 1979). Kratz was thinking of establishing a benchmark that all teachers can be judged against. In his study, "helpfulness" was labelled as the most significant characteristic of a great teacher and "personal appearance" was reported as the next important feature. This can be accepted if one just considers the students apart from other methods when evaluating teachers. That idea was not accepted by Barr (1948) who claimed that supervisors' assessment of teachers was the actual choice metric. However, some researchers started to examine student achievement and use students' learning to infer about teacher quality assuming that supervisors' opinions of teachers do not reveal anything about students' learning. For instance, Domas and Tiedeman's (1950) review of more than 1000 studies of teacher characteristics indicated that for evaluators, there is no clear direction. The notion of using students' achievement to evaluate their teachers was, however, rejected by Getzels and Jackson (1963) who argue that many of the tests were inappropriate to address the effectiveness of teachers. Medley, Coker and Soar (1984) support this opinion claiming that students' achievement may vary and achievement tests can be poor measures of the success of the students themselves. This is true especially because students' achievement can be linked to a wide range of distinct considerations.

   The era of Observing Teaching Performance focused on detecting effective teachers' behaviours that cause student learning. Brophy and Good (1986) argue that learners who receive quality instruction by their teachers achieve more than those who work independently or receive poor instruction. Clark and Peterson (1986) do not only concur with this view but also go further claiming that good teachers tend to adapt their instructions to their students' needs. However, Powell and Beard (1984) argue that subjective judgment can be found when comparing one domain in teacher performance to another. Their bibliography of teacher evaluation research between 1965 and 1980 remains a valuable reference. From the time when it was first commenced until recently, teacher evaluation based on teacher performance has gone through different changes and many concerns have been detected "including evaluation inflation, highly subjective instruments, and a lack of objective measures" (Nagel, 2012, p. 33). Noticeably, the previous overview reflects that despite the fact that there are many methods to assess the quality of teachers; each one has its own limitations. Notwithstanding the restraining factors, the fact may remain that better student learning can be a result of effective instruction (Darling-Hammond, 2000).

3   Why Conduct Evaluation of Teachers?

According to McGreal (1983), evaluation is expected to serve two fundamental needs: Accountability (summative evaluation) and improvement (formative evaluation). The push for both accountability and improvement has resulted in supervision relying on integrated models of formative and summative evaluation (Gullat & Ballard, 1998, p. 16). However, both purposes of teacher evaluation cannot be satisfied by only one system (Towe, 2012). If one system is claimed to satisfy both purposes, one of them is expected to have more weight than the other. Danielson and McGreal (2000) argue that formative evaluation is conducted with the importance placed on teacher improvement, growth, and development. In line with this, Bailey (2007) argues that formative evaluation is conducted mainly to offer feedback or for the purpose of improvement. It might be claimed, then, that formative evaluation can be used to feed professional development decisions. Peterson (2000) supports this and claims that formative assessment data may be used as feedback to shape performances, build new practices or alter existing practices. Summative evaluation, on the other hand, is the summary of evaluation that serves decision-making. Its focus is on ranking, rating, and making judgments about the adequacy of teachers' performance (Danielson & McGreal, 2000). Bailey (2007) argues that the results of summative evaluation help to determine if the funding is going to be continued. Summative evaluation to her is "a final assessment, a make-or-break decision at the end of a project or funding period" (p. 184). However, teachers are not often directly involved in this kind of evaluation.

According to Daresh (2001), a diagnostic evaluation can be considered as a third purpose for teacher evaluation. According to him, this type of evaluation is used to "determine the beginning status or condition (…) prior to the application or intervention or treatment" (p. 281). As such, Bailey (2007) argues that before any attempt to change and in order to provide data about the current status, diagnostic evaluation can be carried out. She also claims that it seems sensible to start with a diagnostic evaluation, followed by systematic formative evaluation, and then a summative evaluation can be conducted after an extended period of formative evaluation. As a sequence, this seems to be logically adequate, however, all three types can be given a different amount of attention and significance depending on the context, objectives, and the rationale of the evaluation system adopted in the educational institution.

4   How to Evaluate Teachers?

In the wide range of literature on teacher evaluation, there have been various methods to evaluate teachers, such as student ratings, peer observation, self-evaluation, and teaching portfolio. In the following section, we will present some of them. They will be presented randomly so that the order does not indicate

priority or significance of one of them over the others. Yet, it depends on the educational institution's needs and characteristics to adopt one or more of them to satisfy the purpose or purposes of EFL teacher evaluation in that particular institution.

4.1 Student Ratings

Student ratings are commonly used to evaluate the performance of teachers. Seldin (2006) argues that it is expected that everyone thinks the ratings of students are all that we need to evaluate the effectiveness of teaching. It may be widely known that students as the product of educational systems have a very close and extended interaction with their teachers; hence, their judgment can be valuable and genuine. Despite the fact that students are seen as a significant source to evaluate the performance of a teacher in a wide range of educational institutes, their ratings as a tool have their own limitations.

Most of the students might not be well prepared nor have enough experience that enables them to evaluate their teachers. Accordingly, they might concentrate on the teacher's personality and give it more attention than academic and teaching skills. Arreola (2007) argues that students in compulsory maths and science courses tend to rate teachers harshly. In line with this argument, students might tend to evaluate EFL teachers harshly when English language is compulsory. In such a case, student ratings can be more beneficial for professional development programs. Accordingly, inclusive evaluation systems will need to consider research findings before counting on student ratings solely. In their study on Japanese university students rating of teaching, Burden and Troudi (2010) support this view and argue that other evaluative methods, such as self-evaluation could be introduced in order to encourage more professional development input.

4.2 Peer Observation

Peer observation can be a useful tool to reflect on the performance of teachers inside their classrooms. It can be more precise, objective, professional and effective than student ratings to develop the instructional practice at educational institutions. Teachers may make use of checklists and forms for peer review that are provided in Braskamp and Ory (1994), Chism (1999), and Weimer, Parrett, and Kerns (2002). Seldin (2006), however, suggests three phases for peer observation; pre-visit consultation where visitor reviews the syllabus and other relevant materials, the visit itself where the visitor observes the performance of the teacher, and the follow-up visit where both of them discuss ideas and observations.

Arguably, serious weaknesses might be highly related to peer observation as a method to evaluate teachers' performance. For instance, how can one make sure

that the piece of teaching that is being observed is representative of everyday practice? Another drawback could be the presence of the observer him/herself that can affect or even alter the class environment and disturb learning. In an attempt to help address previous disadvantages, Arreola (2007) argues that scheduling multiple visits, training peer observer teams, preparing the students, preparing the instructors, and scheduling a post-observation conference might be useful.

4.3   Self-evaluation

Though not widely used as a method for teacher evaluation in many educational systems, self-evaluation can be a good method to evaluate teachers. Teachers themselves perceive the lack of self-evaluation as a weakness in any teacher evaluation system (Towe, 2012). Self-reflection can be significant, since teachers are able to analyse their own instructional practices, which will help towards their professional growth. A major criticism of this evaluative method might be that teachers tend to give themselves higher ratings than they deserve. Besides, this method cannot be used in decisions like promotion (Centra, 1980). Brandt (2010) partly supports this argument and claims that "self-evaluation is a formative, not a summative, activity" (p. 208). This might be true, yet teachers can be more aware than anyone else about their own contributions and hence might be better able than others to annually report their own progress.

4.4   Teaching Portfolio

Teaching portfolios can be used as a means to collect materials and to provide evidence and documents showing the teaching effectiveness of the teacher. Portfolios may also reflect the individuality of teaching. Seldin (2006) argues that "developing a teaching portfolio allows the faculty member to connect theory with practice" (p. 114), which provides the teacher with "a natural outcome of improvement" (p. 114). The key problem with this approach is that it depends on how teachers present their work in the portfolio; accordingly, a very high trust level is required between teachers and principals (Arreola, 2007). In fact, teachers should be trusted especially in reporting and documenting their own work for appraisal purposes in order to have better and more effective teaching. It might be argued that portfolios have more advantages over observation since they represent larger accounts of teaching, yet they might be seen as difficult to deal with from an evaluators' point of view (Alwan, 2010). Accordingly, there needs to be clear criteria and standards to construct portfolios effectively.

To conclude this section on the methods of teacher evaluation, a comprehensive inclusive teacher evaluation system in any educational system needs to consider all the above-mentioned methods along with others (if needed depending on the

purpose of the evaluation) in order to have an adequate evaluative tool. Multiple sources of teacher evaluation techniques can be very useful for principals and administrators of educational institutes to evaluate, improve, and enhance the effectiveness of their teachers. Consequently, a well-designed teacher evaluation system is expected to identify the features of effective teaching and to allocate their effective teaching criterion and accordingly develop the outcome of the whole educational institute. When taking EFL teacher evaluation into account, special concerns may arise for both evaluators and teachers who are being evaluated. The following section will highlight some of the major issues related to EFL teacher evaluation and special attention will be directed to the context of higher education.

5   EFL Teacher Evaluation

In their conceptual articles, Brown and Crumpler (2013) claim that there is no agreement as to what makes any assessment method effective. This problematic issue affects the teachers of foreign languages in particular. They argue that foreign language teacher evaluation has more challenges, especially for the evaluators who do not have sufficient knowledge about second language acquisition and the case becomes worse when those evaluators do not speak the target language that is used inside the classroom. Despite the fact that in this case it is challenging to judge the content knowledge of the teacher and the degree of students' understanding, principals very frequently observe teachers' performance in foreign language classrooms using checklists that contain the content knowledge of the foreign language teacher as one of the criteria for teacher performance assessment. For this particular reason, Brown and Crumpler (2013) call for a change in foreign language teacher evaluation.

Brown and Crumpler (2013) developed a model that positions assessment of peers at top priority of foreign languages instructors' evaluation to shift evaluation towards more learning and progression. Their assessment portfolio model, in Fig. 1, offers an inclusive and wide-ranging instructor's performance assessment that is informed by "multiple sources of evidence, which leads to a more complete and authentic evaluation" (p. 145). They also argue that due to their busy schedules, administrators cannot supervise and evaluate foreign language teachers properly. In fact, their model can be seen as adequate in contexts where self-assessment, as a method for teacher evaluation, is marginalized since this model overlooks the self-evaluation of a foreign language teacher where the teacher him/herself diagnoses his/her teaching in an attempt to improve the quality of his/her own performance.

In an attempt to investigate the main criteria of in-service English language teachers' evaluation, Akbari and Yazdanmehr (2011) conducted an exploratory study in five private language institutes in Iran. Interviews with the supervisors along with analysis of application forms, observation sheets and other relevant documents illuminated the procedures and criteria of teacher assessment in the
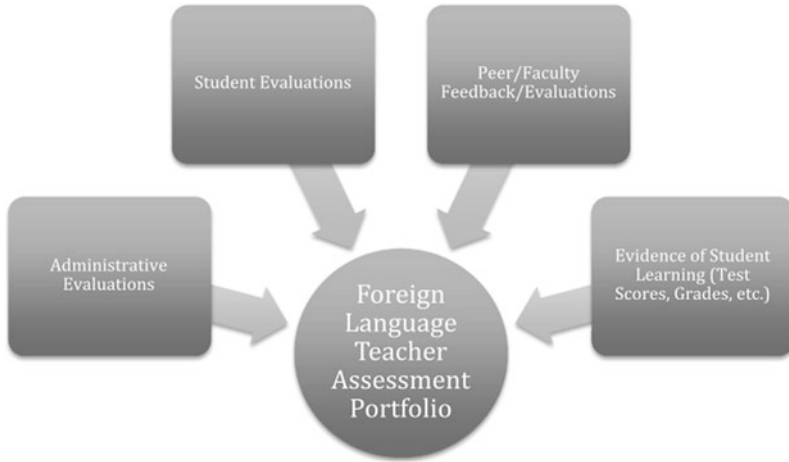
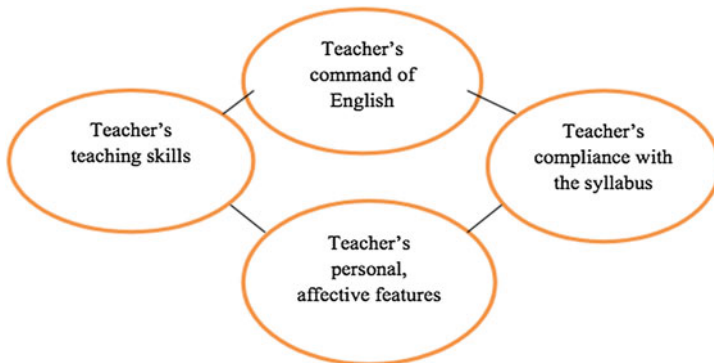**Fig. 1** Brown and Crumpler's (2013) model

Fig. 2 Akbari and Yazdanmehr's (2011) model

target setting. Their procedures in assessing in-service English language teachers' performance are categorized into four groups of teacher's command of English, teaching skills, compliance with the syllabus and personal/affective features. The model they developed exclusively for English language teachers is presented in Fig. 2.

In Akbari and Yazdanmehr's (2011) model, the teacher's command of English involves: Accuracy of speech, structure, pronunciation, and performance in discourse along with fluency in speech. Personal/affective features include: Punctuality, rapport with learners, tolerance in error treatment, enthusiasm and dynamism in involving learners. Teacher's compliance with the syllabus comprises: Expected content to be covered, educational goals to be achieved, and the way to present the material to be followed. Teaching skills involve: Communication skills,

classroom management techniques and task management. Their model might be seen as distinctive and uniquely designed for EFL teachers, however, it does not consider other social and administrative skills besides community service and research activities that might be essential parts of EFL faculty members' activities that should not be overlooked.

By surveying 457 post-secondary foreign language teachers, Bell (2005) examines teacher perceptions on the teaching attitudes and behaviours contributing to effective foreign language teaching. Her study demonstrated a strong positive agreement on all five standards for foreign language teaching. Other categories that teachers agree with the majority of items include: Qualifications of teachers, general theories related to the communicative approach to foreign language teaching, the significance of small group activities, and negotiation of meaning and strategies in foreign language classes. In fact, the study is more concerned with the teachers' behaviour and attitude towards aspects highly related to language acquisition rather than contributing to the effectiveness of foreign language teachers and teaching.

Brown (2006) investigates students and teachers' perceptions of effective teaching in foreign language classrooms that, he argues, are distinctive from other subjects. The findings are the result of analysing a questionnaire distributed amongst 49 university teachers and 1400 of their students. From the teachers' perspective, engaging students in information gap activities, assessing group tasks, being as knowledgeable about culture as language, and having students respond to physical commands are the main characteristics of effective foreign language teachers. Concerning students' opinions, correcting oral errors indirectly, being as knowledgeable about culture as language, having students respond to physical commands, addressing errors with immediate explanation, presenting grammar with real-world context, speaking with native-like control of language, using real-life materials in teaching language culture, and engaging students in information gap activities are the most prominent features of effective foreign language teaching. Arguably, Brown's study can be seen as much concerned with instructional practices and disregarded the other areas that can be used to evaluate language teachers.

Al-Hammad (2011) conducted a study aiming at examining the teaching performance level of 18 English language teachers from the intermediate-level schools in the city of Hail, Saudi Arabia, according to the teaching quality standards. By employing an analytical and descriptive approach, Al-Hammad utilized a controlled observation method on teaching standards and found that the use of teaching aids, and class management skills were highly achieved. In that study and within her sample, students' assessment and lesson delivery were satisfactorily accomplished. Lesson planning was, however, the lowest quality standard achieved by her participants. Al-Hammad's study reinforces the importance of conducting in-service training sessions on teaching quality standards for English language teachers mainly in the three dimensions of: Planning, implementation and assessment. Despite the fact that the sample was solely English language teachers, the dimension and the criteria were not subject-content oriented and could be applicable to teachers of any other subjects. In the Gulf context Al Mahrooqi, Denman, Al-Siyabi, and Al-Maamari (2015) compared Omani school students and teachers'

perceptions of the characteristics of good EFL teachers. One hundred and seventy-one Omani students and 233 English teachers took part in the study which showed general agreement between students and teachers about the importance of all characteristic categories, with special importance to English language proficiency and equality in treating students.

6 Summary and Implications

In this chapter, we have shown how EFL teacher evaluation might be different from other types of teacher evaluation. Purposes and methods may look the same for all subject teachers; however, when conducted for language teachers, evaluators need to consider the uniqueness and the nature of the subject matter. Different concerns for EFL teacher evaluation have been discussed separately and need to be taken into account before determining and constructing academic systems to evaluate language teachers.

References

Akbari, R., & Yazdanmehr, E. (2011). EFL teachers' recruitment and dynamic assessment in private language institutes of Iran. *Journal of English Language Teaching and Learning, 8*, 29–51.

Al-Hammad, Y. (2011). *Teaching performance assessment for the intermediate school English teachers in the city of Hail as per the Teaching Quality Standards.* Master's Thesis, Imam Muhammad Bin Saud Islamic University, Saudi Arabia.

Al-Mahrooqi, R., Denman, C., Al-Siyabi, J., & Al-Maamari, F. (2015). Characteristics of a good EFL teacher: Omani EFL teacher and student perspectives. *SAGE Open*, pp. 1–15.

Alwan, F. (2010). Teacher voice in teacher evaluation: Teaching portfolios in the United Arab Emirates. In C. Coombe, M. Al-Hamly, P. Davidson, & S. Troudi (Eds.), *Evaluating teacher effectiveness in ESL/EFL contexts* (pp. 232–244). Michigan: The University of Michigan Press.

Arreola, R. (2007). *Developing a comprehensive faculty evaluation system: A guide to designing, building, and operating large-scale faculty evaluation systems.* San Francisco: Jossey-Bass.

Bailey, K. (2007). *Language teacher supervision: A case-based approach.* Cambridge: Cambridge University Press.

Barr, A. (1948). The measurement and prediction of teaching efficiency: A summary of investigations. *Journal of Experimental Education, 16*(4), 203–283.

Bell, T. (2005). Behaviours and attitudes of effective foreign language teacher: Results of a questionnaire study. *Foreign Language Annals, 38*(2), 259–270.

Brandt, C. (2010). Giving reflection a voice: A strategy for self-evaluation and assessment in TESOL teacher preparation. In C. Coombe, M. Al-Hamly, P. Davidson, & S. Troudi (Eds.), *Evaluating teacher effectiveness in ESL/EFL contexts* (pp. 199–212). Michigan: The University of Michigan Press.

Braskamp, L., & Ory, J. (1994). *Assessing faculty work: Enhancing individual and institutional performance.* San Francisco: Jossey-Bass.

Brophy, J., & Good, T. (1986). Teacher behaviour and student achievement. In M. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 328–375). New York: Macmillan.

Brown, A. (2006). *Students' and teachers' perceptions of effective teaching in the foreign language classroom: A comparison of ideals and ratings.* Doctoral Dissertation, Graduate College, University of Arizona.

Brown, I., & Crumpler, T. (2013). Assessment of foreign language teachers: A model for shifting evaluation toward growth and learning. *The High School Journal, 96*(2), 138–151.

Burden, P., & Troudi, S. (2010). An evaluation of student ratings of teaching in a Japanese university context. In C. Coombe, M. Al-Hamly, P. Davidson, & S. Troudi (Eds.), *Evaluating teacher effectiveness in ESL/EFL contexts* (pp. 152–166). Michigan: The University of Michigan Press.

Centra, J. (1980). The how and why of evaluating teaching. *Engineering Education, 71*, 205–210.

Chism, N. (1999). *Peer review of teaching: A sourcebook*. Bolton: Anker.

Clark, C., & Peterson, P. (1986). Teachers' thought process. In M. Wittrock (Ed.), *Handbook of research on teaching* (pp. 255–296). New York: Macmillan.

Danielson, C. (2001). *New trends in teacher evaluation: Educational leadership*. Alexandria, VA: Association of Supervision and Curriculum Development.

Danielson, C., & McGreal, T. (2000). *Teacher evaluation to enhance professional practice*. Alexandria, VA: Association of Supervision and Curriculum Development.

Daresh, J. (2001). *Supervision as proactive leadership*. Prospect Heights, IL: Waveland Press.

Darling-Hammond, L. (2000). Teacher quality and student assessment: Are view of state policy evidence. *Educational Policy Analysis Archives, 5*(1), 1–44. doi:http://dx.doi.org/10.14507/epaa.v8n1.2000

Domas, S., & Tiedeman, D. (1950). Teacher competence: An annotated bibliography. *Journal of Experimental Education, 19*(99), 101–218.

Getzels, J., & Jackson, P. (1963). The teacher's personality. In N. L. Gage (Ed.), *Handbook of research on teaching*. Chicago: Rand McNally.

Glasman, N. (1986). *Evaluation-based leadership: School administration in contemporary perspective*. Albany: State University of New York.

Guba, E., & Lincoln, Y. (1981). *Effective evaluation: Improving the usefulness of evaluation results through responsive and naturalistic approaches*. San Francisco: Jossey-Bass Inc.

Gullatt, D., & Ballard, L. (1998). Choosing the right process for teacher evaluation. *American Secondary Education, 26*(3), 13–17.

McGreal, T. (1983). *Successful teacher evaluations*. Alexandria, VA: Association of Supervision and Curriculum Development.

Medley, D. (1979). The effectiveness of teachers. In P. Peterson & H. Walberg (Eds.), *Research on teaching concepts, findings and implications* (pp. 11–27). California: McCutchan Publishing Corporation.

Medley, D., Coker, H., & Soar, R. (1984). *Measurement-based evaluation of teacher performance: An empirical approach*. New York: Longman.

Nagel, C. (2012). *Teachers' perceptions regarding portfolio-based components of teacher evaluation*. Dissertation. Illinois State University.

Norris, N. (1990). *Understanding educational evaluation*. London: Kogan Page Ltd.

Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed.). London: Sage Publication Inc.

Peterson, K. (2000). *Teacher evaluation: A comprehensive guide to new directions and practices*. Thousand Oaks, CA: Corwin.

Powell, M., & Beard, J. (1984). *Teacher effectiveness: An annotated bibliography and guide to research*. New York: Garland.

Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: A systematic approach* (7th ed.). London: Sage Publication Inc.

Seldin, P. (2006). *Evaluating faculty performance: A practical guide to assessing teaching, research, and service*. San Francisco: Jossey-Bass.

Towe, P. (2012). *An investigation of the role of a teacher evaluation system and its influence on teacher practice and professional growth in four urban high schools*. Dissertation. Seton Hall University.

Tyler, R. (1950). *Basic principles of curriculum and instruction*. Chicago: University of Chicago Press.

Verma, G., & Malick, K. (1999). *Researching education: Perspectives and techniques*. London: Falmer Press.

Weimer, M., Parrett, J., & Kerns, M.-M. (2002). *How am I teaching? Forms and activities for acquiring instructional input*. Madison: Atwood.