REVIEW ARTICLE

The Evaluation of Teacher Performance in Higher Education

La evaluación del desempeño docente en la educación superior

Luis F. Gómez *
ITESO, Universidad Jesuita de Guadalajara, Guadalajara, México
ORCID: http://orcid.org/0000-0001-6988-0163

María G. Valdés
ITESO, Universidad Jesuita de Guadalajara, Guadalajara, México
ORCID: https://orcid.org/0000-0001-8802-4995

*Correspondence

Email: lgomez@iteso.mx

Cite as:

Gómez L F., & Valdés, M. G. (2019). The Evaluation of Teacher Performance in Higher Education. *Propósitos y Representaciones,* *7*(2), 479-515. doi: http://dx.doi.org/10.20511/pyr2019.v7n2.255

Summary

The purpose of this article is to present a critical review of the ways in which the performance of university teachers are usually evaluated. Generally, the institutions state proposal of evaluation that usually coincide with what is evaluated and the use of the results. It includes the review of the different evaluation models to show the discrepancies between the objectives and the practices, that is, the poor relationship between the model used and what is done. Universities usually declare that the focus of the evaluation is improvement of practice. One of the most used instruments is the questionnaire, with which an accurate measurement can be made, but generally it does not evaluate the teaching performance, but, in the best-case scenario, the opinion of the students. If universities state that the evaluation should serve to improve the teaching practice, it would be advisable to favor the constructivist model that promotes changes required through the reflection that arises from the review of the self-evaluation, the hetero-evaluation and the co-evaluation.

Keywords: Teaching Performance; Teacher Evaluation Models; Evaluation for Improvement; Higher Education.

Resumen

El objetivo de este trabajo es presentar una revisión crítica de las maneras en que se tiende a evaluar el desempeño de los docentes universitarios. Por lo general, las instituciones enuncian propósitos de la evaluación que suelen no coincidir con lo que se evalúa y con el uso que se hace de los resultados. Se incluye la revisión de distintos modelos de evaluación Para mostrar los desfases entre objetivos y prácticas, es decir, la pobre relación de lo que se hace, con el modelo del declarado. Se tiende a enunciar que el objetivo de la evaluación está centrado en la mejora continua. Uno de los instrumentos más utilizados es el cuestionario, con el que se pretende hacer una evaluación precisa, pero que, por lo general, no evalúa el desempeño docente, sino, en el mejor de los casos, la opinión de los estudiantes. Si se declara que la evaluación debe servir para mejorar la práctica docente, convendría privilegiar el modelo constructivista que permitiría favorecer los cambios que se requieren a través de la reflexión que surja de la revisión de la autoevaluación, la hetero-evaluación y la co-evaluación.

Palabras clave: Desempeño docente; Modelos de evaluación docente; Evaluación para la mejora; Instituciones de educación superior.

Introduction

In reviewing what higher education institutions (HEIs) state about teacher performance evaluation, it is found that they favor teacher learning over practice in order to improve it. However, this statement is not necessarily reflected in the practices since there is a tendency to use the outcomes for other purposes. This action has been much questioned, both in its intentions (undeclared) and in its methods. The former tend to focus on administrative matters and not on academic performance improvement matters, and the latter focus on a one-sided and quantitative view since they focus on student opinion through a single instrument.

In order to achieve consistency between saying and doing, it is necessary to have a clear understanding of what, why, for what and how, which means knowing the different evaluation models. Analyzing each of these elements offers ideas to support the decisions made by the HEIs regarding the actions that best suit the declared intentions. Analyzing the evaluation established in the HEIs is a necessary task to create proposals for the professional development of teachers,

and, therefore, for education quality improvement. This requires to examine specifically the articles, reflections, and debates that have been generated by specialists in the field.

This article is intended to present a critical review of how university teacher performance tends to be evaluated and, based on this, to propose constructivism as a framework consistent with the aim of continuous improvement of teaching practice.

The analysis arises from the exploration of the evaluation process of a HEI where inconsistency between goals, method, instruments, and use of outcomes was detected, and from there, how teacher performance is evaluated in other institutions was examined, and similar difficulties were found. Having found that it is a generalized problem, the existing evaluation models were examined in order to find if any of them offered ways to solve the inconsistencies. In this search, constructivism was envisaged as a perspective from which to propose a way of evaluation that allows teachers and institutions to learn and, therefore, to aim to continuous improvement processes.

Renewing the way of seeing teacher evaluation requires an approach to the different proposals on the matter, and, with it, a clear picture of the advances, challenges, dilemmas, and difficulties that have been experienced in this field of knowledge. The text integrates key elements that provide necessary information for the design and planning of the teacher evaluation process.

The first part of the article presents three of the evaluation models most commonly used in educational institutions, and describes their characteristics, advantages, and disadvantages. This review emphasizes the different meanings of evaluation, as well as the purposes and uses that the educational institutions make of this process.

Given that most HEIs use to a greater extent the questionnaire to obtain data about the dimensions associated with the work of the university teacher, the second part includes information regarding the use given to this evaluation instrument. The emphasis of the third section is placed on the negative judgments that scholars have posed regarding the use of the questionnaire as they consider that its application is based on a technical-instrumental rationality, which is oriented toward the achievement of outcomes, especially those that favor efficiency and effectiveness, and, therefore, reward- or punishment-based control.

The fourth and fifth paragraphs present alternative ways of evaluation, as well as the proposal to use the constructivist framework to place evaluation on the path toward continuous improvement. It concludes with the final reflections on this process.

Evaluation Models and Proposals

In order to show an overview of the theoretical, conceptual and methodological variety of evaluation processes in general, the proposals are grouped into three large models proposed by Rama (1989), Nevo (1983), and Gaytán (2012), based on their analysis and systematization of the different views that have existed regarding evaluation.

Firstly, they mention the models that use goals as the fundamental criterion of evaluation, and that will be called goal-oriented model. One of the authors who stand out in this model is Tyler (Mora, 2004), who considers that the model should be based on the goals set, that is, comparing goals and outcomes in order to determine to what extent the former has been achieved. Tyler's proposal that is framed within the quantitative paradigm focuses on the evaluation of achievements rather than on other variables of the process. According to Pérez (2007), this model is based on the premise that effectiveness is achieved by accomplishing the established goals, which are the only source of the criteria used to evaluate.

Among the negative judgments made of this model, extemporaneousness stands out because the information arrives too late to be used in the teacher improvement or continuous education since this type of evaluation focuses on evaluating the final achievement of the course, and the information it provides has little scope to form a solid basis for issuing a value judgment (Mora, 2004).

In addition to the proposals that are framed in the goal-oriented model, there are those that envisage the evaluation centered on the decisions they serve, regardless of the final outcomes, which will be called decision-making oriented model. In this model, the data obtained is provided to the decision-making bodies. A characteristic feature of this type of model is the continuous and systematic evaluation, which is used to plan and carry out services or proposals aimed at satisfying needs. From this view, evaluation leaves behind the fact of being understood as a control and measurement instrument as it is intended to make an assessment at the end of a process, and comes to be seen as a means for feedback of the educational process (Morales, 2001, p. 169). The proposals of Context, Input, Process, and Product (C.I.P.P) of Stufflebeam and of the University of California Center for Study of Evaluation (C.E.S) of Alkin, and of the Process Evaluation of Cronbach pertain to this model.

The third model, called naturalistic, is based on negotiation and comprehension. It recognizes the important role played by the audience, and the transactional and phenomenological relationship between the participants and the evaluator. Emphasis is placed on the qualitative description of the phenomenon and, to a lesser extent, on the quantitative one. Consequently, the main objective focuses on description and interpretation and not on the assessment and prediction of other models. The parties concerned raise and clarify a number of issues which, in turn, help to identify the aspects and procedures for achieving the desired outcomes (Stufflebeam & Shinkfield, 1995, p. 313). Among the proposals that stand out in this model are Stake's Responsible Evaluation, MacDonald's Democratic Evaluation, Parlett and Hamilton's Illuminative Evaluation, and Eisner's Art Criticism Evaluation.

Each of these models favors a view of evaluation. A variety of conceptions can be recognized in these models: evaluation as the process that allows determining the degree of consistency between the achievements and the previously established goals (Tyler); evaluation as the process of determining, obtaining and providing relevant information that helps decision making (Stufflebeam); evaluation as the process of estimating the merit of something in relation to the intrinsic characteristics of what is being evaluated in order to point out the differences that evaluative judgments would imply in both the use and application in a given context (Scriven & Cronbach, 1982), among others.

Each of the above models favors a criterion that will be used in assessing the information obtained, for example, goal definition is the basic criterion of the models inspired in Tyler's proposal; goal definition in a political context is the fundamental criterion in decision-making models, and standards correspond to the criterion taken into account in the conceptions that emphasize merit as a goal.

From these views of evaluation, it can be seen that the outcome-oriented model tends to be used in the evaluation of teacher performance in HEIs. It is based on the assumption that the student, as the recipient of the service, can evaluate the performance of the teacher, and therefore, based on his opinion, the work of the teacher is evaluated. From this point of view, aspects of the process and of the context are no longer considered. Díaz Barriga (1987) thinks that models that use the goal criterion - observable behaviors - only show their presence or absence without making an effort to understand the meaning of a behavior or to analyze the causes for their presence. In addition to this criticism, Díaz Barriga (1987) points out that the models centered on the provision of information for decision making complicate and hinder the function of evaluations with the goal of determining the sense with which the educational task is projected. This is why he categorizes them as pragmatic and utilitarian, while the outcomes serve decisions.

The choice of teaching evaluation indicators is a problematic issue because, although it is clear that they need to meet several criteria, including: validity, reliability, relevance, timeliness, coherence, clarity, transparency, and accessibility, Chalmers (2007) considers that universities tend to choose performance indicators to include in surveys or questionnaires, precisely because they can be easily quantified and not because they accurately evaluate teachers' performance.

The plurality of conceptions shows that it is possible to evaluate in order to know, assess, and learn, but also to master, justify decisions or achieve goals oriented to immediate levels of reality. Mora (2004) points out the meaning and the uses of evaluation are understood according to the needs, purposes, and goals preferred by a given educational institution. For some, emphasis may be on control and measurement, while for others it may be on the judgment of the validity of the goal or on accountability. Although the concept of evaluation changes depending on the model preferred by each institution, it can be said that the definitions share the idea that it consists in determining merit (Scriven, 1967), that it is a process of issuing judgments (Suchman, 1967), that it is based on established criteria (De la Orden, 1982), and that it requires empirical information (Pérez Gómez, 1983).

Stroebe (2016) points out that, initially, teacher performance evaluation was intended to helping teachers to improve their teaching, and later, it has been used to make important decisions on staff hiring and dismissal, promotions and salary determination; in any case, it will have to be considered whether evaluation is used as a way to better influence learning (formative evaluation) or as a way to rate, reward, or punish performance (summative evaluation).

Rueda and Díaz Barriga (2004) consider that summative evaluation is linked to decision making to decide promotions, salaries, types of hiring, among others. What is relevant in this type of evaluation is the amount of achievements and products, and its purpose is defined through quality standards based on efficiency. Formative evaluation focuses on teacher improvement, continuous education, and professional growth. It is oriented toward studying the aspects of practice that can be improved. Consequently, it provides specific guidance for teacher improvement (Amaranti, 2017. P.97-98). According to Amaranti (2017), it is clear that it is necessary to move forward in the building of an evaluative culture because, despite the importance that has been given to the formative evaluation of teacher performance in HEIs, instrumental rationality continues to be preferred, which is linked more to administrative purposes than teaching quality improvement (p.94).

The questionnaire as a preferred instrument in teacher evaluation

Evaluation practices of HEIs are closer to Tyler's goal-oriented model, and, therefore, tend to use quantitative questionnaires from student perception, often as the only instrument. Thus, it is necessary to analyze this practice and show the inconsistency with the intentions declared by the HEIs.

From 1920 onwards, some universities began to ask students to evaluate teacher performance. Student opinion surveys have been applied in Mexican universities, since the 1960s in the Universidad Iberoamericana (Ibero-American University), since the 1970s in the Accounting School of the Universidad Nacional Autónoma de México (National Autonomous University of Mexico - UNAM), and since the 1980s in the Universidad Autónoma Metropolitana (Metropolitan Autonomous University - UAM) and in the Universidad Autónoma de Aguascalientes (Autonomous University of Aguascalientes - UAA). Since the 1990s, they became widespread among higher education institutions (Canales, Luna Díaz-Barriga, Monroy, Díaz & García 2004).

Studies on the evaluation of university teacher performance around the world show that student opinions surveys are the most used method to evaluate it (Wellein, Ragucci & Lapointe,

2009; Montoya, Arbesú, Contreras & Conzuelo, 2014; Rueda, 2008; Tirado, Miranda and Sánchez, 2007) despite the fact that there are several proposals to evaluate more comprehensively and that probably the only thing measured by these surveys is the course preference, and not the assistance provided by the teacher to help student learning (Stroebe, 2016).

Given this interest in student opinion on teacher performance, databases contain a large number of studies that have been conducted on the validity of questionnaires, item grouping, respondent students' priorities, etc. To cite just a few examples, Durán-García and Durán-Aponte (2015) developed an explanatory model with seven reagents grouped into two factors called attitude and organization. Cortés, Campos, and Moreno (2014) conducted a study on the prioritization of the dimensions of teacher performance evaluation by students. The results showed that the students prioritized the following dimensions: outcomes, learning scenarios, relationships, and evaluation. Luna and Arámburo (2013) studied the relationship between the evaluation scores assigned by students to teacher efficiency and the variables associated with the characteristics of the teacher and the course. Although there are a large number of studies on the validity of questionnaires, it continues to be questioned, but especially its results as the only source of data for evaluation.

The questionnaires ask about the teacher's mastery of the content of the subject, his performance with the group, the use of pedagogical techniques, and the method he uses to evaluate learning, among others (Rueda, 2008). In this long tradition of using questionnaires to evaluate teacher performance, the manner in which they are filled out has changed, but their content and use continue without many changes. In the past, they were filled out manually, now they are often filled out online. In Latin America, there is the same tendency to use questionnaires to measure university teacher performance. In a study conducted by Montoya, Arbesú, Contreras, and Conzuelo in universities in Mexico, Chile and Colombia in 2014, they found that most of the evaluation is carried out through instruments that collect students' opinions.

Student opinion surveys are widely used in Mexican universities, as in Latin America and around the world. According to Rueda (2011), these instruments are used in 80% of Mexican public universities with a large number of irregularities in their design, application process, and result communication. In a study conducted by Rueda (2008) in five major public universities in the metropolitan area of Mexico City, it was found that teacher evaluation is carried out mainly through questionnaires applied by the immediate chief or teacher self-evaluation. The researcher points out that teacher performance evaluation in public universities is driven by salary compensation programs such as PROMEP, PIFI, and others developed by a particular university. The outcomes of the evaluation are used to provide feedback to the teacher and to inform the administrative bodies.

Advocates of the use of questionnaires to evaluate teacher performance present reasons such as those of Molero and Carrascosa (2005) when citing Gillmore:

- It allows students to give their opinions about their teachers, generating political benefits for the institution.
- Students, by their nature, are considered to be the most extensive observers of teaching, and, in turn, are in a unique position to assess the quality of the course, and the dedication and preparation of the faculty.
- The reliability of student observations is usually high. This depends, among other factors, on the number of respondents, so if the appropriate samples are selected, reliability can be assured (p.23).

Gómez and Gaviria (2004) consider that teacher evaluation by students is one of the best systems to evaluate teacher performance because students are the only ones capable of evaluating whether teachers have helped them learn. Despite this view, they argue that questionnaires cannot be the only source of information for teacher evaluation. Tirado, Miranda, and Sánchez (2007)

advocate that students' opinions should be used more rather than being discredited or stigmatized as they believe that students have the right to give their opinion about the quality of the educational services they receive. In the opinion of Tirado, Miranda, and Sánchez (2007), the use of questionnaires has contributed to identifying the characteristics of good teachers, and among the examples, they mention a 2004 study conducted by García in which teachers and students choose the most important characteristics of teaching. The characteristics chosen by teachers were: (1) mastery of the subject, (2) structuring of objectives and contents, (3) class organization, (4) learning evaluation, (4) interaction qualities, and (5) presentation quality (p.15).

The students ordered the characteristics differently. For them, the order of relevance was as follows: (1) mastery of the subject, (2) presentation quality, (3) class organization, (4) interaction qualities, (5) learning evaluation, and (6) structuring of objectives and contents (p. 15).

For these authors, the problem is not in the study of the characteristics of good teachers, but in how teachers can develop these characteristics.

Criticism of the use of questionnaires for evaluating teacher performance

Criticism of teacher evaluation focuses on the validity of the instruments, the purposes of the evaluation, the narrowness of the evaluation, and the timing of the evaluation. It has already been pointed out that evaluation, among its activities, compares performance against a standard. Some of the questions are: What is the standard? Who determines it? Why is it the valid standard? With respect to the method, the questions, among others, are: How is teacher performance measured? What are the instruments used to measure it? What is the validity of its instruments? At what point in the course?

A fundamental question about questionnaires and scales is: Do they really measure teacher performance? Serrano and Arámburo (2013) affirm that the scores assigned to university teachers by students refer to one of the products of the teaching activity: student satisfaction with teaching. Although the questionnaires do not offer a direct measure of the main products of the teaching activity such as learning, alteration of beliefs, etc., they argue that it presents an indirect measure of most products of the teaching activity because they make judgments about the extent to which they consider the teacher affects teaching, in a different sense.

Many researchers, including Alterio and Pérez, (2009), Loaiza, Soria, and Bellido, (2016) and Cortés, Campos, and Moreno (2014.) criticize that it is often not recognized that scales and questionnaires measure only opinions, and that it is not possible to evaluate teaching activity or teacher performance through them. Although it is true that the simplest way to evaluate teacher performance is through the use of questionnaires because they are easy to apply, easy to process through computer programs, and can be displayed as a graph, so as to make comparisons with the scores of others, and, furthermore, that administrators can easily review numerical data, it is also clear that an instrument of this type does not measure teacher performance, it only measures student opinions. From this perspective, that is, if it does not measure what it should measure (validity), the sense of precision and objectivity provided by numbers is meaningless, as are the comparisons of the outcomes. In this regard, Canales et.al (2004) point out that there is no immutable or unique metric in the scales used, and, consequently, the comparison of teachers that is usually made is risky. If teacher "A" obtains an overall average of 5.0 in the scale, teacher "B" an overall average of 2.5, and the faculty an average of 3.5, it cannot be said that teacher "A" is twice as effective as "B". This reading is misleading because it is not possible to control all the variables that can be misleading and bias the data obtained through the instrument (p.126).).

In relation to ethics, Canales et al (2004) use ideas suggested by Sproule (2000) to question student responsibility, in the sense that they answer the questionnaires anonymously, which can cause them to fill them out negligently, falsely, and even maliciously. In addition, they

are rarely asked to argue or substantiate their opinions. Similarly, there are teachers who pressure students to evaluate them in a positive way, conditioning grades or appealing to emotion. These unethical attitudes seriously compromise the validity and reliability of the evaluation due to falseness and bias (p.123).

According to Stroebe (2016), there is empirical evidence to assume that indulgent teachers, that is, those who give their students good grades for little work, are more likely to receive, in turn, good scores from them. On the contrary, on occasion, teachers whose courses ask the student to make a greater effort tend to receive worse scores, so there are serious doubts about the validity of the evaluations students make of their teachers as a measure of the teaching quality or the effort of their teachers (Braga, Paccagnella & Pellizzari, 2014).

Carrel and West (2010) conducted a study finding that students evaluated better the teachers who helped them to learn better the contents of the subject they were studying, which led them to obtain better grades, than those teachers who promoted a deep knowledge that led them to perform better in the subsequent courses, but making a greater effort in the current course, which suggests that students are not evaluating the support they receive for long-term learning.

According to Montoya, Arbesú, Contreras, and Conzuelo (2014), "The validity of the questionnaires in each institutional environment continues to represent an important issue to address. Furthermore, no progress has been made in the differentiation of teaching modalities, or in the nature of the various fields of study in the questionnaires. These noticeable differences are not present in the evaluating instruments" (p. 23).

The evaluation that only considers student opinions is extremely limited, and it is not very useful when such student opinions are collected near the end of the semester because teaching is an interactive process between a group of students and the teacher, and, to a large extent, includes the building of bridges between the meanings of one and the meanings of the other, so that the actions that are perceived as promoters of learning in one group will not necessarily be perceived in the same way by the other. Among the oppositions made to this type of teacher performance evaluation, the following stand out:

- Wolfer and Johnson (2003) identify time of instrument application as an important criticism because, if applied at the end of the semester, teachers find out student opinion upon completion of the course. This limits the possibilities of improvement, and it is likely that one group's evaluation would differ markedly from the other group's evaluation because teaching does not take place in a cultural, ideological, political, and personal affection gap.
- Rueda, Luna, García, and Loredo (2010) find that teaching evaluation has been associated with control and forms of salary compensation to the detriment of the improvement of teacher performance, despite the fact that teaching represents the most important substantive function developed by the University.
- Hénard and Leprince-Riguet (2010) consider that the most controversial method to evaluate education quality is the use of questionnaires by students. The risk with questionnaires is that they can lead teachers to adopt short-term strategies that are harmful to the learning process.
- Canales et al (2004) point out that the questionnaires for teacher evaluation by students are partially based on a learning theory. Therefore, the result-based feedback to the teacher is limited because they do not explore in depth the aspects required to guide him (p.101).
- DeCosta, Bergquist, Holbeck, and Greenberger (2016) refer to the narrowness of evaluation because questionnaires or scales are used as the only source of teacher performance evaluation data.

- Parada (2016) points out that an important limitation of quantitative evaluation systems, as in the case of questionnaires, is the reduction of evaluation to instrumentalization and result, but not to the process of development or transformation of the teaching task, which causes a feeling of unease in the teacher being evaluated, a deep-rooted stress of being approved, and many times to respond to administrative requirements, but not to pedagogical demands (p.185).

Proposals and alternatives for improvement of teacher performance evaluation

Given that all universities have among their goals to improve student learning, teacher evaluation will be valid only to the extent that it contributes to the achievement of the educational purposes and the improvement of teaching practices, and not when it is used as a means of control. As Stiggins (2005) points out, if examination should help students succeed, then why is this same role not played in teacher evaluation?

According to Duro (2015), continuous education improvement requires new ways of thinking about evaluation to come into play because a good system of teacher performance evaluation needs to be based on a clear and consensual definition of the concept of education quality, to have a theoretical framework that coherently explains the system, to determine the purposes clearly and explicitly, to make explicit the consequences of the evaluation, to show consistency between policies, curricula, and conditions of practices, and to include continuous improvement in the purposes.

Although it is known that the objective of teacher performance evaluation should be to provide feedback to teachers in order to influence quality improvement and, at the same time, act as an accountability mechanism, it has often been seen that this is not achieved. This dual-purpose suggests the need for differentiated information when interpreting the data obtained from the evaluation process, that is, to provide administrators with the summary they need regarding teachers' ability to teach and to provide teachers with information that helps them improve teaching by specifically discriminating areas of strength and opportunity. In this regard, Canales et al (2004) emphasize the inappropriateness of making evaluations for control and feedback purposes using the same procedure and instrument (p119).

A well-built evaluation system should help teachers identify strengths and areas of opportunity in their pedagogical practices (DeCosta, Bergquist, Holbeck, & Greenberger, 2016), have elements of judgment different from their own, which will allow them to recognize successes, omissions, and errors, and thus able to improve their practice and generate a higher quality teaching practice (Tirado, Miranda & Sánchez, 2007), and have clarity in their progress toward specific objectives and goals.

Furthermore, Wellein, Ragucci, and Lapointe (2009) argue that different sources of data and instruments are needed for evaluation to achieve a comprehensive view of how University teaches. They propose self-evaluation by the teacher, hetero-evaluation by specialists, students and colleagues, since they both provide information from different perspectives, thus obtaining a more holistic view of teacher performance. Comprehensive evaluation mechanisms not only improve teachers' individual teaching practices, but are also the first steps toward reflective teaching (MacMillan, Mitchell & Manarin, 2010).

With respect to self-evaluation, the teacher can explain the reasons for the course design, the logic of the inclusion of the planned activities, the adequacy of the content, the validity of the sources of information, and the consistency of the evaluation system with the purpose of the course. Likewise, the teacher can explain the reasons for the activities he carried out, the things that helped and hindered his work of supporting student learning, and the strengths and weaknesses he perceives in his performance as a teacher.

Hetero-evaluation can be formed by students' opinions about their learning, the fulfillment of the course goals, and how the activities and resources of the subject helped them or not learn what was proposed in the course program. The products generated by teachers and students can be a very important element to evaluate teacher performance. With respect to co-evaluation, the academy or the faculty coordinator, together with the teacher, can review whether the objective of the course is clearly stated and achievable. Based on the objective, it is possible to review the resources planned as means to achieve it, the teaching method, the content, the evaluation system, the sources of information, the resources used for learning, examples of tasks elaborated by the students, the grades earned by them, and their opinions about their learning and how the course design and the activities implemented by the teacher promoted or not their learning.

Based on a study on teacher evaluation in universities in Mexico, Chile, and Colombia, Montoya, Arbesú, Contreras, and Conzuelo (2014) suggest that the evaluation include: a) the use of diverse instruments and varied sources of information, as well as the use of qualitative methods, b) a link or relationship with the process of teacher training and professional development of teachers, c) an evaluation of the contribution of the collegiate bodies, and not only the individual teacher for a particular course, and d) information in a quick system, easily accessible and readable by teachers. In the same vein, Tirado, Miranda, and Sánchez (2007) point out that, in order for teacher evaluation to be effective, it requires to be viewed in a holistic and integrated manner, and not as an isolated event in the educational organization. In other words, they refer to a systemic process that covers other components, such as the directive function and the study programs, given that each component affects the institutional operation.

In a recent study on the topic, Uttl, White, and Gonzalez (2016), who conducted a meta-analysis of studies on the topic, affirm that the correlation between teacher evaluation and student learning is zero, so they suggest universities begin to give appropriate weight to these evaluations when making decisions about the faculty. In this regard, since 1977 McKeachi has proposed that students evaluate the achievement of educational goals instead of evaluating the teacher's behaviors. Involving students in the evaluation of their learning promotes reflection by students and their teachers, which would give important feedback on the learning process. The issue would be for students to evaluate to what extent they met the learning goals, and not teachers' specific behaviors.

In 2003, Wolfer and Johnson proposed that, if questionnaires are to be applied, students should make only direct judgments about the effectiveness of teaching: extraordinary, appropriate, and unacceptable. They cite a variety of research studies that validate the fact that students make these direct differentiations since the apparent accuracy of the numerical data on the instruments can be misinterpreted as having a level of measurement accuracy that simply does not exist. If student opinion is to influence teacher performance, it needs to be collected, at least, at two points in time: at the middle of the semester and at the end of it. Hallinger (2010), in a study carried out to improve teaching quality, asked students to answer 5 questions (3 closed questions, and 2 open questions) at the middle of the semester. The closed questions, using a Likert scale, focused on: class preparation by the teacher, his ability to involve students, and the general evaluation of the course up to that point. The open questions were: what I like the most about this class is ..., and this class would be better if... This allowed teachers to benefit from students' feedback, and, with it, they could modify their course when it was still time.

In sum, it can be said that an evaluation proposal with intentions other than "control" has to analyze multiple variables related to practice, that is to say, with the meanings, intentions, potentialities, and needs of the agents involved. This means going beyond the comparison between the collected data and the pre-established norms, while the situations recognized as complex are analyzed more deeply. In this regard, Diego and Rueda (2012) point out the following:

A different evaluation, which seeks to include the principal actors of the teaching and learning process, should consider the assessment of other aspects beyond the administrative ones. An evaluation that is more interested in issues such as learning, dialogue, understanding, diagnosis, reflection, and improvement will have greater formative potential than that in which the ultimate aim is grades, comparison, hierarchization, promotion or sanction (p. 62).

In order for an evaluation to fulfill its task, it is necessary to use several sources of data and not just one, thus promoting collaboration, reflection, and inquiry by teachers (DeCosta, Bergquist, Holbeck & Greenberger, 2016). In the same vein, Madaus and Kellaghan (2000) suggest the possibility of combining the data obtained through 1) the application of performance tests that are used in the goal-oriented evaluation model, 2) the resources allocated, considered within the decision-oriented model, and 3) the data from observations and interviews that are typically applied in the naturalistic models. The authors affirm that the combination of models is the best way to document the complexity of the education systems and programs.

According to Stake, quoted by Gaytan (2012), teacher evaluation has to respond to the specific problems of the environment where the teacher performs, understanding that teaching is a complex activity that needs to be analyzed from different perspectives. In addition, teacher participation and institutional environment where the evaluation takes place are indispensable for this process (Amaranti, 2017, p.98).

Teacher performance evaluation not only requires to be comprehensive and used for valid purposes, but also be careful not to damage the teacher's self-image, integrity, and personal and professional dignity (Rueda and Rodríguez, 1996). In the same vein, Tejedor (2012) points out that, in general, how the information provided by non-formative evaluation is produced and used causes in teachers unpleasant feelings, including mistrust, fear, insecurity, and sometimes panic, precisely because institutions tend to use it as a synonym of arbitrariness, irrationality, and authoritarian and overwhelming power (p.321).

The contribution of constructivism to teacher performance evaluation

The constructivist perspective provides principles related to evaluation models focused on processes and transactions, thus highlighting the central ideas that can contribute to the design of a teacher performance evaluation system consistent with the purposes declared by the HEIs, that is, learning and continuous improvement. This approach is also consistent with the suggestions made in the previous section.

How should teachers be evaluated when the pedagogical paradigm declared by a higher education institution is constructivist in nature? Would it consider teachers as permanent learners in search of continuous improvement of their professional practice? Would it consider that the adequate evaluation would be a portfolio with evidence that includes the course program, the materials used, the activities carried out, and the products developed by its students? Would it take into account teacher self-evaluation and peer co-evaluation and student feedback? Or would it take students' answers in a questionnaire as its main input for only improving client satisfaction?

The evaluation that is conceived as a means to achieve the continuous and permanent development of the teacher as an education professional implies to be considered as a fair process consistent with the educational model preferred by the HEIs, which, at the discourse level, tends to be constructivist in nature. In this way, there will be inputs that will not only help evaluate teacher performance objectively, deeply, and impartially, but will also allow teachers to use them to reflect on and improve their practice, which, in the words of Tejedor (2012), will lead to professional advancement based on learning processes.

For his part, Rigo (2008) considers that when the constructivist view is favored in teacher evaluation, the teacher is encouraged to reflect on how to achieve constructivist-oriented teaching,

but also to do so in a constructivist manner (p.132). This implies that this education professional becomes a learner or a constructive specialist who puts into practice what he fosters in his students. From this point of view, evaluation is considered as an inherent part of the learning process and an activity to continue learning. This same author proposes that the dynamics of evaluation be carried out through different activities: comprehensive, explanatory, critical, propositional and interventional.

In what Rigo (2008) calls comprehensive activity, the teacher identifies in a critical and reflective manner the expectations, motivations, beliefs, and assumptions that are present in his performance. In the explanatory activity, he explains the exogenous factors that have a direct or indirect influence on the work with students. In the critical activity, he makes an analytical and substantiated assessment of the context, inputs, processes, and products associated with his task. In the propositional activity, there are proposals and actions that address the deficiencies and areas of opportunity detected, and how strengths will be consolidated. And, in the interventional activity, the actions considered in a previous phase should be implemented. This cycle of identification, action, and evaluation is restarted successively and permanently (p.132).

The evaluation process with the respective phases proposed by Rigo (2008) goes far beyond the ways traditionally used by the HEIs, that is, those that focus evaluation on isolated cognitions that tend not to consider their connection with the general and personal knowledge framework of teachers. When the evaluation is intended to provide teachers with an opportunity to continue learning, it is necessary to consider the broader context involved in the teaching and learning processes. The data provided by evaluation are best understood in the context of a larger system of meanings.

In a sense very similar to Rigo's (2008) proposal, Valliant (2008, p.14) shares the teacher performance evaluation model used in Australia, which includes four stages developed over a two-year period: preliminary planning, data collection, interview, and follow-up. In a meeting at the beginning of a school period, student and teacher review the planning in the manner of a dialogue. In that meeting, they establish the focus of the evaluation and the data sources that will be used to collect the information, which is analyzed and used to establish professional development plans. In the second year, a follow-up is carried out to analyze the support received by the teacher and the progress made with respect to the goals previously established. In this experience, the specification and determination of standards are key because they provide parameters with respect to teaching quality.

In this same vein, Amaranti (2017) points out that teacher evaluation results require to be used as a means to reflect on teaching, and, therefore, to improve it. This necessarily implies that the formative intentionality of evaluation considers feedback and a support system for teachers to have time to introduce the necessary changes in their practices that arise from the analysis and self-evaluation of their work (p. 94). A definition of self-evaluation, modified from Andrade and Du's (2007) proposal, would say that it is a process by which the teacher reflects on and evaluates the quality of his work, and judges the degree of the fulfillment of the goals. Each of these questions would be the criteria for identifying strengths and weaknesses in his work, and, therefore, the criteria for improvement.

Thus, self-evaluation becomes a key factor as it helps the teacher not only become aware of what he does, but also assume the responsibility of reflecting on his practice critically and propositionally in order to improve it. García, Loredo, Luna Rigo, and Rueda (2011) consider that teacher evaluation should consider, at least, four fundamental components: teacher thought, teaching and learning planning, teaching practice in the classroom, and evaluation of student results (p.30). These components can serve as a basis for analysis and reflection.

Given that the constructivist approach is interested in promoting that the learner (teacher) has deep knowledge, focuses on real problems, and improves his metacognition and reflective skills, ways of evaluation have been designed, including portfolios, self-evaluation, peer

evaluation, and co-evaluation with the teacher (Yurdabakan, 2011). In this regard, Canales, Luna, Díaz-Barriga, Monroy, Díaz, and García (2004) point out the existence of alternative methods that have been developed in recent dates to evaluate teaching, among them, standing out case analysis, writing of diaries with critical incidents, analysis of teacher discourse, dilemma approach and resolution, building of cognitive maps that represent teachers' epistemological beliefs (p. 134). For his part, Tejedor (2012, p.323) recommends the use of self-reporting as one of the instruments for university teachers to provide information on their teaching activity. With this instrument, they can recognize the problems they face with respect to planning, plan development, and evaluation, and identify how they solve problems, thus having the possibility of providing important judgments and reflections to improve their teaching practice. The use of these means of evaluation constitutes an important aid for teachers to realize and explain the set of implicit theories, beliefs, and values that are present in everyday practice.

Crispin and Marvan (1999) highlight the use of the teacher portfolio, which is a systematic and organized collection of products that show the learner's efforts, development and achievement in specific areas. With this instrument, there is the opportunity to analyze, systematize and reflect on their everyday activity. The content of the teacher portfolio represents an opportunity for the teacher to do self-evaluation exercises according to a set of established criteria, but also co-evaluation and hetero-evaluation exercises. There are opposing positions with respect to co-evaluation, among them, that it is time-consuming, and that the evaluation is made from the subjectivity of the evaluators. The first criticism is valid; the second is not. Since teacher performance cannot be measured in grams, meters or seconds, it certainly depends on the subjectivity of the evaluator. However, the use of criteria for evaluation helps objectivity to some extent. Evaluation through opinion questionnaires is not more objective because it counts frequencies, because in that case what is in question is the validity of the instrument.

Peer co-evaluation of teacher performance is related to the idea that teachers form a community of practice and that through their participation they understand and articulate its values and standards (Wenger, 1999). The purpose of the evaluation is for teachers to compare their self-perception of teaching with the evaluation of their students and peers, bringing a critical reflection that will make them aware of the strengths and weaknesses of their teaching practice (Amaranti, 2017, p.97). Since it is clear that, from a constructivist point of view, learning is conceived as a product that derives from the active position of the teacher and the student, it is that the participation of the latter in the evaluation of the teacher is pointed out as another important element that will allow the creation of what Vygotsky (1979) calls "zone of proximal development", a learning space that is also shaped from the view and experience of the student with respect to the practice and performance of the teacher. The student as creator of meaning can contribute his view using instruments that allow evidencing the forms of joint activity between teachers and students. The task then is to redirect the object of the evaluation to the learning activities, the means or the educational aids the teacher offers to the student to build knowledge.

From this perspective, it can be said that teacher performance evaluation will necessarily have to consider a deep knowledge of the teacher, which can only be obtained if the evaluation processes take into consideration the teacher's needs, interests, previous knowledge, learning styles, teaching styles, motivations, work habits and routines, attitudes and values, difficulties, among other aspects. This means that evaluation needs to be seen as a systematic, continuous and formative activity as it contributes to teacher permanent learning, giving an opportunity to improve the formative processes. The evaluation shown as a strategy of professional growth should provide elements for the university teacher to reflect on his practice systematically, deeply and critically, a reflection that not only provides a diagnostic view, but also delves into the trajectories of specific teaching procedures of the profession (Álvarez, 2008, p. 104).

Specifically, it is proposed that teacher performance be evaluated through three processes: self-evaluation, hetero-evaluation, and co-evaluation. In self-evaluation, the teacher self-evaluates the design of the course, the logic of the inclusion of the activities, the adequacy of the

content, the consistency of the evaluation system with the purpose of the course, the things that helped and hindered his work of supporting student learning.

Hetero-evaluation takes into account student opinions, at the middle of the semester and at the end of it, about their learning, the fulfillment of the course goals, and how the activities and resources helped them to learn what was proposed in the program.

In co-evaluation, the teacher, together with the academy or the faculty coordinator, will review, before the course, if the objective is clearly stated and if it is achievable. Based on this, it is reviewed if the resources and the planned method are appropriate. At the end of the course, the self-evaluation and the hetero-evaluation are reviewed in a critical and propositional manner to improve the teacher's practice.

Conclusions

Teacher evaluation has had a strong influence on the quantitative paradigm proposed by Tyler in the 1950s. From this perspective, teacher performance evaluation places emphasis on the verification and measurement of results, which are obtained mostly from student opinion in the questionnaires applied at the end of a given school period. Being a summative evaluation oriented to the fulfillment of objectives or effectiveness, it is related to control and sanction by the administration. García, Loredo, Pérez, Reyes, Rigo, and Rueda (2011) consider that the evaluation that is associated with results and that is administered through instruments and scales to measure discrete behaviors corresponds to an evaluation approach that is oriented to the establishment of significant relationships between the behavior of the teacher and the indicators of the teaching product that are defined as the level of student achievement or the satisfaction or opinion they report. Control-based evaluations have little to do with learning, continuous education, and improvement that are declared at discursive level in the educational institutions, a situation that leads to a problem because the purposes of evaluation that are socialized do not coincide with what is sought or with the uses of the results.

Teacher performance evaluation is a pending task in the higher education institutions because their preferred ways of evaluation tend to reductionism in the sense that they make invisible the political, social, and cultural context that forms part of the reality being evaluated. By focusing mostly on student opinion about certain characteristics being evaluated, it is not known that variables such as initial and permanent training, working conditions, type of subject, curriculum, number of students per group, and type of subject intervene in teacher performance. Teacher intentions and the logic of the course, among other issues, are also ignored. Therefore, it is necessary that the evaluation proposals take into consideration the context in which the teacher performs his work. In this regard, Rueda et al (2014) consider that the educational events take place in specific historical contexts; for the same reason, teachers are immersed in this dimension; for this reason, they insist that the pedagogical must be understood from a social-cultural perspective; otherwise, it would be very difficult to understand the phenomenon of teaching as a whole.

Before making decisions on what to evaluate, it is necessary to have a clear idea of the purpose, since the meaning and value of that activity is in the answer. It is important to explain the purposes of evaluation since they are the basis of the decisions made regarding the theory, the evaluation model, and the design that is most relevant to that purpose. Each purpose comes with what and a how. No approach can give valid and reliable results for any purpose. Teacher performance evaluation needs to clearly distinguish the summative evaluation from the formative evaluation, since one of the great difficulties that have been experienced when implementing this or that evaluation system is the lack of clarity with respect to the purposes.

Having a clear idea of the complexity and multi-dimensionality of the work and the role of the teacher in student formative processes, the evaluation proposals must necessarily open

possibilities to the diverse information sources, and, with it, the opportunity to triangulate said information. By focusing on a single view, that is, that of students, there is a risk of disregarding the complexity involved in the comprehensive study of teaching (Canales, et al, 2004). By opening possibilities to broaden the views, including that of the person involved, good results are obtained, among them, the teacher is placed as an active agent and not as a passive subject of evaluation.

In order for the evaluation to be understood as a means of professional learning and to allow teachers to orient and strengthen their activity, it must be considered as a process that requires a precise definition in their starting models, as well as an appropriate systematization (Vaillant, 2008). From this consideration, it is necessary for each HEI to be able to establish and define a teaching framework from which the different actors participating in self-evaluation, co-evaluation, and hetero-evaluation have clear criteria and indicators that facilitate the implementation of the teacher performance evaluation.

References

Alterio Ariola, G., & Pérez Loyo, H. (2009). Evaluación de la función docente según el desempeño de los profesores y la opinión estudiantil. *Revista de Educación Médica Superior.* 23(3), 1-14. Recuperado de: http://www.bvs.sld.cu/revistas/ems/vol23_3_09/ems01309.pdf

Álvarez Méndez, J. (2008). *Evaluar para conocer, examinar para excluir*. Madrid: Morata.

Amaranti, M. (2017). Uso de resultados de la evaluación docente para mejorar la calidad de la docencia universitaria. Recuperado de: www.congresouniversidad.cu/revista/index.php/rcu/article/download/804/759/

Andrade, H., & Du, Y. (2007). Student responses to criteria-referenced self-Assessment. *Assessment and Evaluation in Higher Education*, *32*(2), 159-181. Doi: https://doi.org/10.1080/02602930600801928

Blum, R., & Arter, J. (1996). *Student performance assessment in an era of restructuring.* Alexandria, VA: ASCD.

Braga, M., Paccagnella, M., & Pellizzari, M. (2014). Evaluating students' evaluations of professors. *Economics Of Education Review*, *41*(4), 71-88. Doi: https://doi.org/10.1016/j.econedurev.2014.04.002

Canales, A., Luna, E., Díaz Barriga, F., Monroy, M., Díaz, M., & García, J. M. (2004). Aproximaciones metodológicas al análisis y la evaluación de la docencia. En M. Rueda y F. Díaz Barriga (Coords.), La evaluación de la docencia en la universidad. Perspectivas desde la investigación y la intervención profesional (pp. 87-201). Ciudad de México: CESU-UNAM.

Carrell, S., & West, J. (2010). Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors. *Journal of Political Economy*, *118*(3), 409-432. Doi: https://doi.org/10.1086/653808

Chalmers, D. (2007). *A review of Australian and international quality systems and indicators of learning and teaching*. Carrick Institute for Learning and Teaching in Higher Education Ltd, Sydney, NSW.

Cortés, E., Campos, M., & Moreno, M. P. (2014). Priorización De Las Dimensiones De Evaluación Al Desempeño Docente Por El Estudiante, En Tres Áreas Del Conocimiento. *Formación Universitaria, 7(2),* 3–10. Doi: http://doi.org/10.4067/S0718-50062014000200002

Crispín, M. & Marván, L. (1999). El portafolio como herramienta para mejorar la docencia, en Rueda, M. y Landesman, M. (coords). *¿Hacia una nueva cultura de la evaluación de los académicos?* México: CESU-UNAM, pp. 184-2002.

De Diego, M., & Rueda M. (2012). La evaluación docente en educación superior: uso de instrumentos de autoevaluación, planeación y evaluación por pares. *Voces y Silencios: Revista Latinoamericana de Educación,* 3(2), 59-76. Recuperado de https://revistas.uniandes.edu.co/doi/pdf/10.18175/vys3.2.2012.04

De la Orden, A. (1982). Integración institucional en la formación del profesorado. *Revista de educación*, *30*(269), 121-126. Recuperado de: http://www.educacionyfp.gob.es/dam/jcr:a77231eb-5938-48c3-ac18-a1141ad65a66/re2690713059-pdf.pdf

De Costa, M., Bergquist, E., Holbeck, R., & Greenberger, S. (2016). A Desire for Growth: Online Full-Time Faculty's Perceptions of Evaluation Processes. *Journal of Educators Online*, *13*(2), 19-52. Recuperado de: https://www.learntechlib.org/p/161394/

Díaz Barriga, A (1987). Problemas y retos del campo de la evaluación. *Perfiles Educativos*, 37, 3-15.

Durán-García, M., & Durán-Aponte, E. (2015). Adaptación y validez de un instrumento para la evaluación de docencia universitaria: escala de desempeño docente institucional (EDDI). *Perspectiva Educacional, Formación de Profesores, 54(1),* 75-89. http://doi.org/10.4151/07189729-Vol.54-Iss.1-Art.306

Duro, E (2015). Tendencias y alcances en el campo de la evaluación educativa en América Latina. Complementariedad entre modelos externos y procesos auto evaluativos en pos de la mejora continua. IV Conferencia ReLac. Recuperado de: http://www.unicef.org/argentina/spanish/Edu_IVConferenciaReLacFINAL.pdf.

García, B., Loredo, J., Luna, E., Rigo, M., & Rueda, M. (2011). Algunas consideraciones sobre los aspectos teóricos involucrados en la evaluación de la docencia. En: M. Rueda y F. Díaz-Barriga. *La evaluación de la docencia en la universidad. Perspectivas desde la investigación y la intervención profesional,* México: UNAM/Plaza y Valdés.

Gaytán, S. (2012). Evaluar la docencia: estudio meta-evaluador del desarrollo del proceso de encuestas al alumnado y su fiabilidad en el Grado en Biología de la Universidad de Sevilla. En A. Castro y otros (Coords.), *Calidad, evaluación y encuestas de la docencia universitaria* (41-53). Murcia: Laborum.

Gómez, A., & Sacristán, G. (1983). *La enseñanza: su teoría y su práctica*. Madrid: Akal Editor.

Hallinger, P. H. (2010). Using faculty evaluation to improve teaching quality: A longitudinal case study of higher education in Southeast Asia. *Educational Assessment, Evaluation & Accountability*, *22*(4), 253-274. Doi: http://doi.org/10.1007/s11092-010-9108-9

Hénard, F., & Leprince-Ringuet, S. (2010). *The path to quality teaching in Higher Education*: OECD Publications. Recuperado de: http://www.oecd.org/edu/imhe/44150246.pdf

Loaiza Jara, O., Soria Quijaite, J., & Bellido Mamani, E. (2016). Redes bayesianas para el estudio de la influencia de la evaluación docente en el rendimiento académico de los estudiantes de la Facultad de Ingeniería y Arquitectura de una universidad privada confesional. *Apuntes Universitarios: Revista De Investigación*, *6*(1), 61-77. Recuperado de: https://revistas.upeu.edu.pe/index.php/ra_universitarios/article/view/112

Luna Serrano, E., & Arámburo Vizcarra, V. (2013). Variables asociadas a la competencia docente universitaria en la opinión de los estudiantes. *Archivos Analíticos de Políticas Educativas*, *21* (1). Recuperado de: http://epaa.asu.edu/ojs/article/view/1098

MacMillan, M., Mitchell, M., & Manarin, K. (2010). *Evaluating teaching as the first step to SoTL*. Paper presented at SoTL Commons Conference, Statesboro, GA.

Madaus, G. F. & Kellaghan, T. (2000). Models, metaphores and definitions in evaluation, en D. Stufflebeam, G. F. Madaus y T. Kellaghan. *Evaluation models: viewpoints on educational and human services* (19-32). Boston: Kluwer Academic Publishers.

McKeachie, W. J. (1997). Student ratings: The validity of use. *American Psychologist, 52,* 1218-1225. Recuperado de: http://www.dartmouth.edu/~chance/course/Syllabi/97Dartmouth/day-14/eval-5.pdf

Molero, D., & Carrascosa, Juan (2005). La evaluación de la docencia Universitaria. Dimensiones y variables más relevantes. *Revista de Investigación Educativa, 23*(1), 57-*84.* Recuperado de: https://www.redalyc.org/html/2833/283321951005/

Montoya, J., Arbesú, I., Contreras, G., & Conzuelo, S. (2014). Evaluación de la docencia universitaria en México, Chile y Colombia: Análisis de experiencias. *Revista

Iberoamericana de Evaluación Educativa, 7(2e), 15-42. Recuperado de: http://www.rinace.net/riee/numeros/vol7-num2e/art1.pdf

Mora, A. I (2004). La evaluación educativa: Concepto, períodos y modelos. *Revista Electrónica Actualidades Investigativas en Educación, 4*(2). Doi: http://doi.org/10.15517/AIE.V4I2.9084

Morales, J. (2001). La Evaluación en el Área de Educación Visual y Plástica en la ESO (Tesis doctoral) Universitat Autònoma de Barcelona, España. Recuperado de: https://www.tdx.cat/bitstream/handle/10803/5036/jjma08de16.pdf.PDF

Nevo, D. (1983). The conceptualization of educational evaluation: An analytical review of the literature. *Review of Educational Research, 53*(1), 117-128.Doi: https://doi.org/10.3102/00346543053001117

OCCDE. 2009). *A conceptual framework and examples of country practices*. México: Autor.

Parada Romero, L. B. (2016). Sistema de evaluación docente, instituciones de educación superior tecnológica: lineamientos de calidad. *Praxis & Saber*, *7*(13), 177-198. https://doi.org/10.19053/22160159.4171

Pascual Gómez, I., & Gaviria Soto, J. (2004). El problema de la fiabilidad en la evaluación de la eficacia docente en la universidad: una alternativa metodológica. *Revista Española De Pedagogía*, *62*(229), 359-375. Recuperado de: https://www.jstor.org/stable/23765063

Pérez, D. (2007). *Revisión y análisis del Modelo de Evaluación Orientada en los Objetivos (Ralph Tyler - 1950)*. https://www.researchgate.net/publication/220025586_Revision_y_analisis_del_Modelo _de_Evaluacion_Orientada_en_los_Objetivos_Ralph_Tyler_-_1950

Rama, G. (1989). *Cambio social, educación y crisis en América Latina*. Análisis, N°.49-50.

Rigo, M (2008). Constructivismo educativo, actividad y evaluación del docente: relato de algunas posibles incongruencias. *Revista Reencuentro*, *53,* 125-134. Recuperado de: https://www.redalyc.org/html/340/34005311/

Rueda Beltrán, M., Schmelkes, S., & Díaz-Barriga, A. (2014). La evaluación educativa. Presentación del número especial de Perfiles Educativos 2013. La evaluación en la educación superior", en *Perfiles Educativos, 36*(145), 190-204 bDoi: https://doi.org/10.22201/iisue.24486167e.2014.145.45995

Rueda, M. (Coord.) (2011). *¿Evaluar para controlar o para mejorar?* México: UNAM-IISUE/ Bonilla Artiga Editores.

Rueda, M., & L. Rodríguez (1996). La evaluación de la docencia en el posgrado de la UNAM. En M. Rueda, y J. Nieto (comps.). *La evaluación de la docencia universitaria*. México: Facultad de Psicología-UNAM.

Rueda, M., Luna, E., García, B., & Loredo J. (2010). La evaluación de la docencia en las universidades públicas mexicanas: un diagnóstico para su comprensión y mejora. *Revista Iberoamericana de Evaluación Educativa, 3*(1), 77-92. Recuperado el 20 de enero de 201, de http://www.rinace.net/riee/numeros/vol3-num1_e/art6.pdf

Rueda, M., & Díaz-Barriga, F. (Coord). (2004). *La evaluación de la docencia en la universidad. Perspectivas desde la investigación y la intervención profesional.* México: UNAM, Plaza y Valdés.

Stiggins, R. (2005). From formative assessment to assessment for learning: A path to success in standards-based schools. *The Phi Delta Kappan, 87*(4), 324–328. Doi: https://doi.org/10.1177/003172170508700414

Stroebe, W. (2016). Student evaluations of teaching: no measure for the TEF. Times Higher Education. Recuperado de: https://www.timeshighereducation.com/comment/student-evaluations-teaching-no-measure-tef.

Stufflebeam, D., & Shinkfield, A. J. (1987). *Evaluación sistemática: guía teórica y práctica*. Madrid: Paidós-MEC. Barcelona: Paidós.

Stufflebeam, D., & Shinkfield, A. (1995). *Evaluación sistemática - guía teórica y práctica*. España: Centro de Publicaciones del Ministerio de Educación y Ciencia. Ediciones Paidós Ibérica.

Tejedor, F.J. (2012). Evaluación del desempeño docente. *Revista Iberoamericana de Evaluación Educativa, 5*(1), 318-327. Recuperado de: http://www.rinace.net/riee/numeros/vol5-num1_e/art24.pdf

Tirado Segura, F., Miranda Díaz, A., & Sánchez Moguel, A. (2007). La evaluación como proceso de legitimidad: la opinión de los alumnos. Reporte de una experiencia. *Perfiles educativos*, *29*(118), 07-24. Recuperado de: http://www.scielo.org.mx/pdf/peredu/v29n118/v29n118a2.pdf

Uttl, B., White, C. A., & Gonzalez, D. W. (2016). Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related. *Studies In Educational Evaluation*, (Preprints).

Vaillant, D. (2016). Algunos Marcos Referenciales en la Evaluación del Desempeño Docente. *Revista Iberoamericana de Evaluación Educativa, 1*(2) 7-22. Recuperado de https://revistas.uam.es/index.php/riee/article/view/4663/5100.

Vygotsky, L.S. (1979). *El desarrollo de los procesos psicológicos superiores.* Barcelona: Grijalbo.

Wellein, M. G., Ragucci, K. R., & Lapointe, M. (2009). A peer review process for classroom teaching. *American Journal of Pharmaceutical Education*, *73*(5), 1-7. Doi: https://doi.org/10.5688/aj730579

Wenger, E. (1999). *Communities of Practice: Learning, Meaning, and Identity*. Cambridge: Cambridge University Press.

Wolfer, T. A., & Johnson, M. M. (2003). Re-evaluating student evaluation of teaching: the teaching evaluation form. *Journal of Social Work Education, 39*(1), 111-121. Doi: https://doi.org/10.1080/10437797.2003.10779122

Yurdabakan, İ. (2011). The View of Constructivist Theory on Assessment: Alternative Assessment Methods in Education. *Journal Of Faculty Of Educational Sciences*, *44*(1), 51-54. Recuperado de: http://www.acarindex.com/dosyalar/makale/acarindex-1423869893.pdf