The Evaluation of Teacher Performance in Higher Education

Joseph Morales
Higher University of San Andres (UMSA) La Paz, Bolivia
Email: moralerjoshep43@gmail.com

Abstract

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.

A. INTRODUCTION

When reviewing what higher education institutions (HEIs) declare regarding the evaluation of teaching performance, it is found that they privilege the teacher’s learning over their practice in order to improve it; however, the discourse is not necessarily reflected in the practices since there is a tendency to use the results for other purposes (Rozi et al., 2022). This way of proceeding has been highly questioned, both in its (undeclared) intentions and in its methods (Gomez & Valdes, 2019). Regarding the former, these tend to focus on administrative issues and not on improving academic performance; the latter focus on a unilateral and quantitative vision, since they focus on the student’s opinion through a single instrument (Dunst et al., 2020).

To achieve congruence between saying and doing, it is necessary to be clear about what, why, what for and how, which implies knowledge of the different evaluation models (Gomez & Valdes, 2019). Analyzing these elements in each of them provides ideas to support the decisions made by HEIs regarding the actions that best suit the stated purposes. Analyzing what is done around the evaluation that has been established in HEI Institutions is a necessary task to generate proposals aimed at the professional development of teachers, and therefore to improve educational quality. For this, a specific approach to the productions, reflections and debates carried out by specialists in the field is required.
The intention of this article is to present a critical review of the ways in which the performance of university teachers tends to be evaluated and, based on it, to propose constructivism as a framework consistent with the intention of continuous improvement of teaching practice (Rozi et al., 2022).

The analysis arises from the exploration of the evaluation process of an HEI in which inconsistency was detected between purposes, method, instruments and use of the results, from this the ways in which it is carried out in other institutions were reviewed and found similar difficulties. Finding that it is a widespread problem, the existing evaluation models were consulted in order to find if any of them offered ways to resolve the inconsistencies. In this search, constructivism was glimpsed as a perspective from which to propose a way of evaluation that allows teachers and institutions to learn and, therefore, to be oriented towards continuous improvement processes (Zubaidah et al., 2021).

Renewing the view in the field of teacher evaluation requires an approach to the different proposals on the subject, and thus have a clear picture of the progress, challenges, dilemmas and difficulties that have been experienced in this field of knowledge. The text integrates key elements that offer necessary information for the design and planning of the teacher evaluation process (Fernandez, 2018). In the first part of the article, three of the most used evaluation models in educational institutions are presented; its characteristics, advantages and disadvantages are described (Tirado et al., 2021). From this review, the different meanings of evaluation are emphasized, as well as the uses and purposes that educational institutions make of this process.

Given that in most HEIs the questionnaire is used to a greater extent to obtain data about the dimensions associated with the work of the university professor, in the second part information is included regarding the use that has been given to this instrument of evaluation. The accent of the third section is placed on the criticisms that scholars of the subject make regarding the use of the questionnaire, while they consider that its application obeys a technical-instrumental rationality, which is oriented towards the achievement of results, especially those that privilege efficiency and effectiveness and therefore the control that is based on reward or punishment (Ozgenel & Mert, 2019).

In the fourth and fifth sections, alternative forms of evaluation are presented, as well as the proposal to use the constructivist framework that directs it towards continuous improvement. It concludes with the final reflections on this process.

B. METHOD
This research uses qualitative research with descriptive analysis approach. The type of data used in this study is secondary data obtained from literature studies and the state of education in Bolivia. Presentation of data using a qualitative approach.

C. RESULT AND DISCUSSION
1. Evaluation Models and Proposals
With the purpose of showing an overview regarding the theoretical, conceptual, and methodological range of evaluation processes in general, the proposals grouped into three large models that Rodriguez (2020) propose with based on the analysis and systematization that they carried out of the different visions that have existed with respect to the evaluation.

In the first place, they point to the models that use the objectives as the fundamental criterion of the evaluation and that will be called the objective-oriented model. One of the authors highlighted in this model is Tyler, who considers that it should start from the goals set, that is, it is about comparing objectives and results to determine to what extent the former have been achieved. Tyler’s proposal, which is part of the quantitative paradigm, focuses on the evaluation of achievements, rather than on other process variables. According to Ayllon et al. (2019), this model starts from the premise that effectiveness is achieved from the achievement of the established objectives, which are the only source of the criteria used to evaluate.

Among the criticisms made of this model, the extemporaneity stands out, since the information arrives too late to be used in the improvement or improvement of the teaching activity, since this type of evaluation focuses on assessing the final achievement of the course and the information it provides has little scope to constitute a solid basis when making a value judgment.

In addition to the proposals that are framed in the model by objectives, there are those that visualize the evaluation focused on the decisions they serve, beyond the final results, which will be called a decision-oriented model. In this model, the data obtained is put at the service of the instances that make decisions. A characteristic note 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 point of view, evaluation leaves behind the fact of being understood as an instrument of control and measurement, since it intends to carry out an assessment at the end of a process, and is now seen as a means that enables feedback on the process. educational. The proposals of Context, Input, Process and Product (C.I.P.P) by Stufflebeam and that of the Center of the University of California for the study of Evaluation (C.E.S) by Alkin, and the Process Evaluation by Cronbach correspond to this model.

The third model called naturalistic, is based on negotiation and understanding. This recognizes the important role of the audience and the transactional and phenomenological relationship that exists between the participants and the evaluator; the emphasis is placed on the qualitative description of the phenomenon and, to a lesser extent, on the quantitative one. For this reason, the main objective is focused on the description and interpretation and not on the evaluation and prediction of other models. Among the interested parties, a series of questions are raised and clarified, which in turn help to identify the aspects and procedures that allow the desired results to be achieved (Gallardo, 2020)). Among the proposals that stand out in this model are: the Responsible Stake evaluation, the Democratic evaluation of MacDonald, the Illuminative evaluation of Parlett & Hamilton and the Critical artistic evaluation of Eisner.
Each of these models favors a vision of evaluation. In them, a variety of conceptions can be recognized, among others, evaluation as the process that allows determining the degree of congruence between the achievements and the previously established objectives (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 evaluated, in order to point out the differences that value judgments would entail both in use and in application for a given context (Allen et al., 2020).

Each of the above models privileges a criterion to be used in the evaluation of the information obtained, for example, the definition of objectives constitutes the basic criterion of the models inspired by Tyler's proposal; The definition of objectives situated in a political context is the fundamental criterion in the decision-making models and the standards correspond to the criterion that is taken into account in the conceptions that emphasize merit as an objective.

From these views of evaluation, it can be seen that in the evaluation of the performance of teachers in HEIs there is a tendency to use the results-oriented model. It is assumed that the student, as recipient of the service, can evaluate the teaching performance, therefore, based on her opinion, the teacher's work is evaluated. From this point of view, aspects of the process and the context are no longer seen. Efendi et al. (2021) considers that the models that use the criterion of objectives-observable behaviors - only record their presence or absence without making an effort to try to understand the meaning of a behavior or to analyze the causes why they are present. In addition to this criticism, Pinheiro et al. (2019) points out that the models that focus their value on the provision of information for decision making hinder and hinder the function of evaluations that have as their goal, the determination of the meaning with which the task is projected educational, which is why he classifies them as pragmatic and utilitarian, while the results are put at the service of the decision.

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

The plurality of conceptions shows that it can be evaluated to know, value and learn, but also to dominate, justify decisions or to achieve objectives that are oriented towards immediate planes of reality. Iglesias et al. (2021) points out that the meaning and uses of evaluation are understood according to the needs, purposes and objectives favored by a certain educational institution; For some, the emphasis may be on control and measurement, while for others it may be on judging the validity of the objective or on accountability. Although the concept of evaluation changes according to the model that is favored in each institution, it can be said that the definitions share the idea that it consists of determining merit, that it is a process of issuing judgments,
which is based on established criteria (Darwanstah, 2021) and which requires empirical information.

Stroebe (2016) points out that, initially, the evaluation of teaching performance was intended to help teachers improve their teaching, later it has been used to make important decisions about hiring and firing staff, for promotions and salary determination. In any case, it will have to be considered whether it is evaluation as a way to better influence learning (formative evaluation) or as a way to qualify, reward or punish performance (summative evaluation).

Timang et al. (2021) consider that the 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 number of achievements and products and its purpose is defined by quality standards based on efficiency. The evaluation for training purposes focuses on the improvement, improvement and professional growth of the teacher; It is oriented to investigate the aspects that can be improved in the practice, for the same reason that it gives concrete improvement guidelines to teachers. For Amaranti (2017), the need to continue advancing in the construction of an evaluative culture is clear, because despite the relevance that has been given to the evaluation of teaching performance in HEIs for training purposes, rationality continues to be privileged instrumental, which is linked more to administrative purposes than to improving the quality of teaching (p.94).

2. The Questionnaire as a Privileged Instrument in Teacher Evaluation

The evaluation practices of HEIs are closer to Tyler’s model by objectives and, therefore, tend to use quantitative questionnaires from the perception of students, frequently as the only instrument. Therefore, it is necessary to analyze this practice and show the inconsistency with the intentions declared by the HEIs.

Starting in 1920, some universities began asking students to evaluate the performance of professors. In Mexican universities, surveys have been applied on the opinion of students since the 1960s at the Universidad Iberoamericana; in the Accounting Faculty of the National Autonomous University of Mexico (UNAM) since the seventies; at the Autonomous Metropolitan University (UAM) and at the Autonomous University of Aguascalientes (UAA) since the 1980s. Starting in the 1990s, they began to generalize among higher education institutions (Alalwan et al., 2019).

Studies about the evaluation of the teaching performance of university professors in the world show that the questionnaires on the perception of the students are the most used method to evaluate it (Maritasari et al., 2020) despite the fact that there are several proposals to evaluate in more comprehensive ways and that probably the only thing that the questionnaires measure is the liking for the courses and not the help that provided by the teacher to encourage students to learn (Stroebe, 2016).

Given this interest in the perception of students about the performance of teachers, in the databases there are a large number of studies that have been carried
out on the validity of the questionnaires, the grouping of the items, the priorities of the students who evaluate, etc. To cite just a few examples, Durán-García and Durán-Aponte (2015) developed an explanatory model with seven items grouped into two factors called: attitude and organization. Cortés, Campos and Moreno (2014) carried out a study on the prioritization of the dimensions of evaluation of teaching performance by students. The results showed that the students prioritized the dimensions: results, learning scenarios, relationships and evaluation. Luna and Arámburo (2013) investigated the relationship between the evaluation scores given by students to teaching effectiveness 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 the questionnaires, it continues to be questioned, but especially its results as the only source of data for the evaluation.

In the questionnaires, questions are asked about the mastery that the teacher has of the content of the subject, about his development with the group, the management of pedagogical techniques and the method used for the evaluation of learning, among others (Rueda, 2008). In this long tradition of using questionnaires to assess teacher performance, the way of filling them out has changed, but the content and the use given to them continue without many variations. Before they were filled out manually, now they are often done online. In Latin America, there is the same tendency to use questionnaires as a measure of the teaching performance of university professors. In a study carried out by Montoya, Arbesú, Contreras and Conzuelo at universities in Mexico and Chile and Colombia in 2014, they found that most of the evaluation is carried out through instruments that collect the perception of students.

In Mexican universities, as is the case worldwide and in Latin America, student opinion questionnaires are widely used. According to Hurbut (2018), these instruments are used in 80% of Mexican public universities with a large number of irregularities in their design, application process and communication of results. In a study carried out by Rueda (2008) in five important public universities in the metropolitan area of Mexico City, it was found that teacher evaluation is done mainly through questionnaires applied by the immediate boss or by self-evaluation of teachers. The researcher points out that the impetus for the evaluation of teaching performance in public universities comes from salary compensation programs with programs such as PROMEP, PIFI and others developed by a particular university. The results of the evaluation are used to provide feedback to the teacher and to inform the administrative authorities.

Those who advocate the use of questionnaires to assess teacher performance present reasons such as those of Kaiser (2019) when they cite Gillmore: a) It allows students to give their opinion about their teachers, causing political benefits for the institution; b) It is considered that students, by their nature, are the most extensive observers of teaching, and in turn are in a unique position to assess the quality of the course, the dedication and the preparation of the teaching staff; c) The reliability of the observations made by the students is normally high. This depends, among other
Li & Wang (2021) consider that student evaluation of teachers is one of the best systems for judging teacher performance because teachers are the only ones capable of judging whether teaching has helped them learn. Despite this opinion, they maintain that questionnaires cannot be the only source of information for teacher evaluation. Quezada et al. (2020) advocate that the opinion of students be used more instead of discrediting or stigmatizing it, since they consider that students have the right to express their opinion about the quality of the educational services they receive. In the opinion of Huisman et al. (2019) the use of questionnaires has contributed to identifying the characteristics of good teachers and among the examples they cite a study by García, from 2004, in which teachers and students choose the most important features of the teaching. For the teachers, they were: (1) mastery of the subject, (2) structuring of objectives and content, (3) organization of the class, (4) evaluation of learning, (4) qualities of interaction and (5) expository quality.

The students ordered the traits differently, for them the order of relevance was as follows: (1) mastery of the subject, (2) expository quality, (3) organization of the class, (4) qualities of interaction, (5) evaluation of learning, (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 to get them to develop them.

3. Criticism of the Use of Questionnaires for the Evaluation of Teaching Performance

Criticisms made of teacher evaluation center around the validity of the instruments, the purposes of the evaluation, its narrowness, and the time in which it is carried out. It has already been pointed out that evaluation has among its operations the comparison of performance against a standard. Some of the questions are: what is the standard? who determines it? why is that the valid standard? Regarding the method, it is questioned, among other things, how is teaching performance measured? with what instruments? What is the validity of your instruments? At what point in the course?

A fundamental question regarding questionnaires and scales is, do they really measure teacher performance? Serrano & Arámburo (2013) affirm that the scores that students assign to university teaching 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 teaching such as learning, alteration of beliefs, etc., they argue that it presents an indirect measure of most of the products of the teaching activity because they make judgments about the extent to which they consider that the teacher affects teaching, in a different sense.

Many researchers, including Rivai et al. (2019), and Zhang et al. (2021) criticize that it is often not recognized that what the scales and questionnaires measure are only opinions and the evaluation of the activity or the teaching performance cannot be done through them. Although it is true that the easiest way to evaluate teacher performance...
is through the use of questionnaires, since they are easy to apply, to process through computer programs, which can be graphed and thus establish comparisons with the scores of others and that, in addition, administrators can easily review the numerical data, it is also clear that an instrument of this type does not measure the performance of the teacher, it only measures the opinions of the students. From this perspective, that is, if it does not measure what it should measure (validity), the sense of precision and objectivity provided by the numbers is meaningless, as are the comparisons made with the results. In this regard, Mislia et al. (2021) point out that in the scales that are used there is no immutable or unique metric, therefore the comparison that tends to be made between teachers is risky; If professor "A" obtains a global average of 5.0 on the scale, professor "B", 2.5, and the faculty of professors an average of 3.5, it cannot be said that professor "A" is twice as effective as " B"; this reading is misleading because it is not possible to control all the variables that can confuse and bias the data obtained through the instrument.

In relation to ethics, Canales et al., (2004) take up ideas from Sproule (2000) to question the responsibility of the students, in the sense that they answer the questionnaires anonymously, which can cause them to fill them out in an anonymous way negligent, false and even malicious. In addition to the fact that they are rarely asked to argue or substantiate what they think. Similarly, there are teachers who pressure students to evaluate them positively, conditioning grades or appealing to emotion. These unethical attitudes seriously compromise the validity and reliability of the evaluation due to falsity and bias.

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

Carrel and West (2010) carried out a study in which they found that students better evaluated teachers who helped them better learn the content of the subject they were studying, which led them to obtain better grades, than those teachers who promoted a in-depth knowledge that led them to perform better in subsequent courses, but involved putting in more effort in the current course, suggesting that students are not evaluating the support they receive for long-term learning.

According to Montoya et al. (2014) “The validity of the questionnaires for each institutional context continues to represent an important problem to address. On the other hand, neither has progress been made in distinguishing teaching modalities, or in the nature of the various fields of study in the questionnaires. These notable differences are not present in the evaluation instruments”.

The evaluation that only contemplates the opinion of the students is extremely limited and when it is collected towards the end of the semester it is of little use, given that teaching is an interactive process between a group of students and the professor
and that to a great extent includes build bridges between the meanings of one and the other, so that the actions that are perceived as promoters of learning in one group are not necessarily going to be perceived in the same way by another.

D. CONCLUSION

By being clear about the complexity and multidimensionality that exists in the task and in the teacher’s role in the students’ training processes, the evaluation proposals must necessarily open possibilities to the diversity of information sources and with it the opportunity to triangulate that information. By focusing on a single vision, that is, that of the students, there is a risk of making invisible the complexity involved in the comprehensive study of teaching. By opening possibilities to broaden the views, including that of the person involved, gains are obtained, among them, the teacher is placed as an active agent and not as a passive subject of the evaluation. From this consideration it is necessary that each HEI can outline and define a teaching framework from where the different actors that participate in self-assessment, co-assessment and hetero-assessment have clear criteria and indicators that facilitate the implementation of teacher performance evaluation.

REFERENCES


