Rev. Innova Educ. (2020). Vol 2 No 1 pp. 25-44

Revista Innova Educación

www.revistainnovaeducacion.com

ISSN: 2664-1496 ISSN-L: 2664-1488

Edited by: Instituto Universitario de Innovación Ciencia y Tecnología Inudi Perú

Process of designing instruments for teaching performance evaluation: experience in a distance education unit of higher education

Proceso de diseño de instrumentos para la evaluación del desempeño docente: experiencia en una unidad de educación superior a distancia

Karla Ávila[1]
Universidad Autónoma de Yucatán, Mérida – Yucatán, México
https://orcid.org/0000-0001-5372-9143

Leonor López[2]
Universidad Autónoma de Yucatán, Mérida – Yucatán, México
https://orcid.org/0000-0001-9701-5193

DOI: https://doi.org/10.35622/j.rie.2020.01.002

Received **20/12/2019**/ Accepted **25/01/2020** Published **29/01/2020**

ORIGINAL ARTICLE

KEYWORDS

Distance education, evaluation, teaching performance, university.

The main objective of this study was the design and implementation of instruments to evaluate teacher's performance in a higher education distance unit of a public mexican university. The process consisted of an eight-stage program, based on information generated by interviews, group discussions and document revisions. As a result, three instruments were designed and implemented to collect information from different perspectives: from the student, the supervisor and the evaluated person himself.

PALABRAS CLAVE

Educación a distancia, evaluación, desempeño docente, universidad.

El objetivo de este estudio fue el diseño e implementación de instrumentos de evaluación del desempeño de docentes a nivel licenciatura en la modalidad a distancia, adscritos a una unidad específica para esta modalidad dentro de una universidad pública mexicana. El proceso se llevó a cabo a través de un programa de ocho etapas, el cual fue elaborado con base en información recabada a través de entrevistas, discusiones grupales y revisión documental. Como resultado, se diseñaron y administraron tres instrumentos de evaluación para recoger información desde distintas perspectivas: del estudiante, el supervisor y el mismo evaluado.

[1] Correspondence: karla.avila@ correo.uady.mx
[2] Email: leonor.lopez@correo.uady.mx

This work is under a Creative Commons Attribution 4.0 international license.

1. INTRODUCTION

The e evaluation can be understood in different ways: as a tool to regulate and measure, verify the validity of the objectives or as an accountability mechanism; this perspective will be in function of the needs, goals or purposes of whom applies it (Mora, 2004). To define it, Rosales (2014) based on the definitions by Stufflebeam, Lafourcade and the UNESCO, explains that the evaluation allows to acquire information and to issue opinions about certain phenomena, situations, individuals or objects, providing tools for decision-making.

On the other hand, the concept of job performance is constructed by Chiavenato (2009) as the effectiveness of the personnel to execute their functions inside an organization, whilst Milkovich and Boudrem (as referred to, in Chiang and San Martín, 2015, p. 160) ratify that the performance references how much a worker meets the mandatory requirements for his position. According to the idea that the success of an organization is determined, to a large extent, by their workers, the urgency to know the contribution that they make to the organization was increasing and with that the idea of evaluating the performance, although initially only the employers perception was taken into account, becoming increasingly participatory through the upcoming years (Sánchez and Calderón, 2012).

Several authors have provided their own formalization for the term "performance evaluation". According to Sastre and Aguilar (2003), it is "that systematic and structured process, of monitoring the professional work of the employee, to assess his performance and the results achieved in the performance of his position" (p. 321). That said, Castillo (2007) delineates it as "the process of periodic judgment of the quality of work and labor potential of the members of a Company" (p. 306). Another definition is that coined by Chiavenato (2009, p. 245), who states that "it is a systematic assessment of the performance of each person based on the activities they perform, the goals and results they must achieve, the competencies offered and their development potential".

Having said that, Gómez-Mejía, Barkin, and Cardy (2001) indicate that performance measurement involves identifying, gauging and managing the yield of the people within the organization. There are instruments used to measure performance, which can yield specific judgments (comparing performance of some employees with that of others) or absolute judgments (individual ones, based on standards); in addition, its approach can be the characteristics, behaviors or results of the workers (Gómez-Mejía *et al*, 2001). There are diverse instruments to employ in performance measurement, such as 360º evaluation, where all those who have contact with the individual being evaluated; the rating scales, in which it is evaluated in regard of certain adjectives

This work is under a Creative Commons Attribution 4.0 international license.

with the help of a scale; or labor standards, which allow comparing the performance of a worker with a predefined standard, just to name a few (Mondy and Noe, 2005).

The task of evaluating performance is part of the human resources management processes, which Castillo (2007) defines as an administrative system that includes the planning, organization, coordination, direction and control of essential activities that result in the appropriate conditions that propitiate the correct development of the employees potential, proposing a model that includes the five processes listed, but framed in the external context and resulting in work performance, where the evaluation is part of its control processes. By contrast, Chiavenato (2009) proposes a different vision, because it argues that classical models treat people as resources that need to be managed, standardized and that are uniform and inert; therefore, it proposes a model of dynamic processes, which are impacted by the influences of both external and internal environments. The basic processes presented by Chiavenato (2009) are six: processes to integrate people, to organize, reward, develop, retain and audit them, forming the performance evaluation part of the processes for organizing people.

Performance evaluation can have two purposes: administrative and development. The first focuses on the use of conclusions to make decisions about working conditions, such as rewards or dismissals; while, development-oriented, the evaluation seeks to improve the performance of workers and the strengthening of their skills, so it should include advice aimed at improvement or training plans (Gómez-Mejía *et al*, 2001). The information to evaluate the performance can come from different sources, such as the evaluated one, managers, subordinates, co-workers, a commission for it or a mixture of the previous ones (Chiavenato, 2009).

However, the above considerations can be applied to any type of organization; then again, those that offer educational services have special characteristics that make it subject to a more specialized evaluation, either because of the institutional imperative related to the social function historically assigned to it, which is to train people (Careaga, 2001) or because educational institutions emphasize quality, which is related to concepts such as effectiveness, efficiency, relevance and satisfaction of those involved in the educational phenomenon (Alcón and Esteve, 2017).

One of the major concerns when evaluating in higher education institutions is the work of teachers, who become an important objective as they are considered central actors in the educational process, a "key guidance subject," as asserted by Prieto (2008, p.325). The teacher's role, due to the complexity of teaching and the nature of the activities he carries out, makes the evaluation of teaching performance a necessity and, at the same time, one of the biggest challenges in universities (Fernández and Luna, 2004).

This work is under a Creative Commons Attribution 4.0 international license.

In Mexico, the evaluation of teachers has had a late appearance and progress, because before the 80's it was practically non-existent (Vidal, 2009, cited in Guzmán, 2018); after this date the evaluation took a central role in the search for improvement of educational quality in public institutions, being such a priority that policies were promoted by the Mexican government to evaluate and grant incentives to university professors (Rueda, 2018), although prior to this there was a boom in the evaluation of professors in private universities, which have been carried out for more than 4 decades (Rueda, Elizalde, and Torquemada, 2003). One of the most used techniques are the teaching evaluation questionnaires (Cuestionarios de Evaluación de la Docencia, CEDA in spanish), which allowed the evaluation of teaching performance and student satisfaction. The CEDAs were used for the first time in an institution of higher education in Mexico in 1972 -by the Universidad Iberioamericana- and it's use was propagated to other universities, although its use has aroused much controversy due to the punitive use they're given (García, 2003), while other academics point out that students are not reliable sources to evaluate teachers, since they lack knowledge and experience to do so; yet, CEDAs are the most used instruments not only in Mexico, but worldwide (Izar, Ynzunza and Castillo, 2016).

Pacheco, Ibarra, Iñiguez, Lee, and Sánchez (2018) report that performance evaluation of university teachers is understood as the "systematic exercise that, based on a set of evidence, judges how teachers carry out their work […] In order to identify achievements and mismatches of their performance in educational settings" (p. 1), while Rico, Montalvo, and Ayala (2001) conceive it as a "systematic assessment of the performance, or qualifications of the professionally defined role, in addition, of the ideology of the school system". Both definitions indicate that evaluation must be a systematic act, and for the design of a teacher evaluation system, Mateo (2000) points out four phases: ideation, development, implementation and meta-evaluation. Ideation consists in identifying needs, deriving institutional objectives and defining the teaching quality model. Then, in the development phase the roles and responsibilities of the teachers are determined, the criteria with which they will be evaluated, indicators are developed and standards for the evaluation are set. During the implementation phase the teaching activity is documented, value judgments are issued, and the improvement proposals are derived from the above, to finally analyze the process itself during the meta-evaluation phase.

This effort can have two purposes, very similar to those proposed by Gómez Mejía *et al.* (2011) and which are mentioned in previous paragraphs: the formative and the summative (Salinas, 2017), the first focuses on the improvement of teacher practices, aiming to identify strengths and weaknesses; while the second is directed to the collection of information useful for decision-making of an academic-administrative type, such as compensations or teacher re-recruitment. One

This work is under a Creative Commons Attribution 4.0 international license.

of the problems that Gómez and Valdés (2019) found that occurs frequently in higher education institutions is the lack of congruence between the purposes of evaluation, the methods and instruments used and the use given to the results, since they found that many universities propose (in their narrative) an evaluation of teaching performance directed to improvement, but in its execution it reflects a tendency towards other ends.

The evaluation of teaching performance varies both in the concepts that explain the phenomenon, the methods for collecting information and the uses that are given to it (García, 2003). Pacheco *et al*. (2018) point out two aspects to be taken into account when evaluating university teachers: due to their high specialization, they need to have two types of knowledge: disciplinary and pedagogical, since in addition to their own knowledge of the discipline(s), they need those related to the educational phenomenon and the training of people in learning scenarios. It is also essential to establish the characteristics of a good teacher. Although efforts have been made to define them in a general way, Pacheco *et al*. (2018) highlight that these should be consistent with the institution's ideology, disciplinary conditions, evaluation needs and the use that will be given to the results. In addition to the above, Elizalde and Reyes (2008) suggest that a contextual analysis of the institution should be taken into account.

In distance education, it is far more difficult to evaluate teachers, since there is not even a solid regulatory framework for this modality (Vicar, 2015), in addition to the fact that the teaching role is different for this schooling technique, having tasks such as: planner and designer, content expert, educational pedagogue-technologist, specialist and technician in the development of teaching materials for digital media, responsible for guiding the educational process, tutor/ consultant as well as evaluator, plus, the demand of being at the same level of a teacher who develops in classroom mode (García, 2003), although without mixing the functions and roles for each one, then, as Fainholc (2004) points out, a good classroom teacher is not necessarily a good distance teacher, and as Páez maintains (2010), there is an "error in transferring the principles and strategies used in the face-to-face mode to the virtual form" (p.153).

There are several particularities of a distance modality teacher, including that it is no longer the only or main source of knowledge, and many times it is not even responsible for generating learning activities; in addition to that, communication is usually asynchronous, the attention is personalized and there is no physical contact, adopting the role of motivator, companion and co-evaluator, plus that requires very specific skills to manipulate digital tools (Páez, 2010).

Teaching performance evaluation in virtual environments has two main approaches, the partial one and the global one. The partial approach is specifically oriented towards some of the aspects of importance for virtual training, such as developmental activity, materials, technological

This work is under a Creative Commons Attribution 4.0 international license.

platform or the relationship between cost and benefit obtained. On the other hand, the global approach considers the total set of elements involved in the process, with the aspects of those focused on quality models or standards and the benchmarking ones, where the organization is compared with another (of excellent results) which aspires to emulate (Rubio, 2003).

Currently, a considerable number of authors have developed proposals and models to evaluate teachers in virtual environments, some intending to evaluate the courses, and others focusing exclusively on this first figure. Among those who recommend to evaluate the teachers performance are Llarena and Paparo (2006), who propose to evaluate students satisfaction, as well as teachers regarding the quality of the materials, the tutors performance and the technological environments quality; while Villar (2008) focuses on the materials, the teacher's interactions with the students, his academic background and the use of the platform.

Regarding to the proposals partial focus, the following are mentioned: Duart and Martínez (2001) propose that virtual teachers' performance evaluation should be carried out considering the perspective of the students, the faculty (coordinators and managers of the program) and the results obtained -focusing on teacher knowledge, learning orientation, motivation, evaluation, relationship with the teaching team and clarity and speed in the answers-. Esquivias, Gasca and Martínez (2009) propose an evaluation to teachers through rubrics, where the evaluated and an observer can locate their performance in the areas of metacognitive, psychopedagogical, communicative, technological, ethical or social competence, and at the novice levels, apprentice, practitioner or expert.

On the other hand, Cabero, Llorente, and Morales (2018) present a proposal to configure a model with four dimensions to be evaluated: disciplinary, pedagogical, technological knowledge and rule and regulation compliance, with a series of subdimensions for each area. They propose the use of questionnaires combined with more qualitative methods such as a portfolio, to collect information from the students, principals or academic managers and the evaluated perspective. This proposal is based on that of other authors, being an influence on the evaluation criteria the TPCK model of Mishra and Koehler (2006), which classify the knowledge that a teacher should have in technological, pedagogical and of content, which are not static nor act separately, but combine and interact.

After reviewing the previously presented background, it can be seen that teacher performance evaluation, in addition to being a complex phenomenon whose varieties fluctuate depending on the modality and indicators to be evaluated, becomes a necessity for educational institutions, especially if you are looking for improvement through the analysis of the results found in the evaluation process. Therefore, this was one of the main concerns of the unit in which the

This work is under a Creative Commons Attribution 4.0 international license.

evaluation instruments were designed, which was created in January of 2019 and began with its functions in August of the same year, with an initial offer of two degrees from the social sciences area.

The educators who participated teaching in the first course (called virtual advisors in this unit) have face-to-face teaching experience, but little or none in distance mode. Due to this issue, the coordination of the unit paid primary attention to teacher evaluation, this in order to identify whether teachers are fulfilling the assigned tasks and functions correctly, as well as to identify gaps and areas of improvement in their performance, which can serve as a basis for the improvement of the introduction courses to the unit, as well as those of continuous training, so that the service that the advisors offer to the students is a satisfying one.

The authors of this work were invited to collaborate in the design of instruments of evaluation, acting as consultants and documenting the process. The instruments were designed in conjunction with the unit's collaborators and other experts in the area of distance education and statistics.

The fundamental objective of this endeavor was the elaboration of teaching evaluation instruments for university professors in virtual modality, for later implementation. To achieve this, specific objectives will be to analyze the current state of the unit, identify the tasks and functions of virtual advisors, and the proposal and implementation of a program for the design of instruments for assessing teacher performance.

2. METHODS AND MATERIALS

In order to gather the necessary information to support the proposed program and build the evaluation instruments, qualitative methods were used: review of the organization's own bibliography and documents (such as manuals, regulations, etc.), semi-structured interviews and guided group discussions.

The construction of the items that were included in the questionnaires was carried out using the method proposed by Hernández, Fernández, and Baptista (2014) to operationalize the dimensions contemplated in the selected model, and subsequently the pilot test was carried out in the first administration, in which students participated voluntarily evaluating each teacher with whom they took courses. Statistical tests were performed to calculate the reliability and validity of the items of the questionnaire directed at students.

3. RESULTS

This work is under a Creative Commons Attribution 4.0 international license.

After performing a diagnosis, an eight-step program was projected to implement the design of evaluation instruments, which consisted of:

I. Selection of a teaching performance evaluation model

For this stage, an investigation of theoretical models (described in the introduction) was carried out in addition to the models used by the organization for its face-to-face and mixed programs, so that there is an overview of how the teacher evaluation is carried out and it is possible to make decisions.

After reviewing the bibliographical material, an analysis of the virtual advisor manual was executed, identifying the tasks and functions that correspond to this figure and classifying them according to the courses stage in which they should be implemented, summarized in the following points:

Previously to the class:

- Review the Virtual Classroom (structure, activities, resources) and notify any changes that must be made
- Verify that the students enrolled in the Virtual Classroom are those who have been assigned to the course, comparing with the list provided
- Make a welcome video and introduction to the course

During the class:

- Enable, at the beginning of the week, the unit and activities corresponding to it
- Resolve students' academic and technical questions through forums, messages and email within a period not exceeding 48 business hours.
- Encourage and monitor that interactions in the Virtual Classroom occur in a climate of respect
- Motivate students to complete the activities and continue with the course
- Review, rate and feedback the learning activities carried out by the students
- Maintain constant communication with the tutor to report incidents with students

At the end of the class:

- Review, rate and feedback the final evaluation
- Load the grades in the Virtual Classroom and in the corresponding system (SICEI)
- Clarify doubts, if any, regarding the final grade of the course
- Self-evaluate your performance

General matters:

- Attend to the meetings convoked by the Coordination

This work is under a Creative Commons Attribution 4.0 international license.

- Participate in courses, workshops, congresses and other continuing education activities, indicating attendance and showing evidence to the Coordination to add it to the file

Therefore, it was decided to opt for the evaluation model proposed by Cabero *et al*. (2018), indicated in the introduction, who through different information sources evaluate the performance of distance mode teachers in the disciplinary, pedagogical, technological and normative areas.

One of the sources of information are the students, who directly receive the service of the virtual advisor and can evaluate their performance. To collect the opinions and perceptions of the students, the creation of an evaluation instrument containing reagents that question the student about the performance of the advisor in the different areas proposed by Cabero *et al*. (2018), it's suggested. This instrument will be complemented with the observation made by the advisory supervisor (supported by an observation guide) and the assessor's self-evaluation (through a questionnaire with open response items).

II. Design of the teaching performance evaluation instruments

For the design of the students instrument, the dimensions proposed by Cabero *et al*. (2018) were associated with the tasks and functions of the advisor. In plenary with unit staff, an advisor on statistics and high school personnel from the same online university, who are more experienced in the area of distance education, questions were written for the instruments.

In the development of the items for the students' questionnaire, the base method was the one proposed by Hernández *et al*. (2014). For this, the first three dimensions (disciplinary, pedagogical and technological) were taken into account, given that compliance with standards is difficult for students to observe, and this will be evaluated by the supervisor. Those subdimensions that are related to the action of the virtual advisor were selected, and writing indicators for each one, which must be observable for the students.

Analysis and modifications were made during meetings with the working group, discussing and reflecting the objective of the evaluation, that is, what is sought to be evaluated; as well as the actions to be evaluated can be fully observed by the students.

After several modifications, the students instrument consists of 34 items, where 30 of them are based on a Likert scale, with options from 1 to 4, to measure the degree of agreement or disagreement with the claims about the advisor's actions; and 3 reagents use a Likert scale, which also offers values from 1 to 4, but assessing the level of satisfaction with certain aspects. Finally, one open type question was added, in which the student is expected to give his general opinion regarding the performance and strengths as well as areas of improvement of the assessed advisor.

This work is under a Creative Commons Attribution 4.0 international license.

As a complement to the evaluation carried out by the students, the supervisors direct observation will be taken into account, they will rate the fulfillment of various regulatory aspects with a 1 if assigned and completed the activities, zero if it is an activity that doesn't applies, or -1 if this activity was assigned, but no compliance was seen. One of the advantages when evaluating the advisors in the virtual modality, is that those who supervise have access to their executed actions, so that the monitoring is close. The supervisor keeps a weekly track of the activities performed by each virtual advisor, recording compliance or non-compliance. This responds to the recommendation of Cabero *et al*. (2018), who in addition to the four dimensions -disciplinary, pedagogical and technological knowledge- suggest that compliance with rules and regulations be evaluated, proposing those regarding vocabulary, the student and the frequency of entry to the platform, to which it was added compliance with other specific unit rules.

It is intended to obtain a numerical score through the results in the evaluation carried out by the students and the supervisor, and for this purpose a formula that weighs the scores of each student was designed together with the statistical advisor and the staff in charge of the advisors area, and a percentage corresponding to the approval index in their subject is added:

$$IDD=(.15*AD+.35*AP+.25*AT+.25*AA)$$

In the formula, called the Teaching Performance Index (Índice de Desempeño Docente in spanish, IDD), the areas evaluated in the instrument for students are involved: disciplinary area (AD), pedagogical area (AP) and technological area (AT), as well as what is evaluated by the supervisor, which is the administrative area (AA). The formula does not include self-assessment, as each advisor is expected to use their results to reflect when making a comparison between the score obtained in the performance index and their own perception.

Regarding the evaluation carried out by the different characters, the total is considered according to the importance given to each area, conforming to the opinion of the staff in direct relation with the teachers of the unit, resulting in 15% for the disciplinary area, 35% for the pedagogical area, 25% for the technological area and 25% for administrative aspects. The reason for the pedagogical area to be the most important is that the majority of the advisor's tasks and functions relate to it, since the advisor is basically a facilitator that mediates the student's interaction with the virtual classroom resources to achieve the learning. The next ones in value are the technological area and the normative aspects, the first one due to the relevance of the advisor mastering technological tools (his work is 100% at a distance), and the normative aspects guide the advisor to give a close accompaniment to the student and a quick response, vital to give human meaning to the online mode, where you cannot always directly observe the teacher with whom

This work is under a Creative Commons Attribution 4.0 international license.

you interact. The area with the lowest weighted significance is the disciplinary one, since those in charge indicate that it is necessary that the advisor be an expert in the contents to be imparted, this to enable the correct resolution of doubts by the students, however, this remains in the background to prioritize the other areas.

To complement these points of view, the opinion of the advisor himself will also be collected, this through an academic follow-up questionnaire, where in addition to being asked for data on the results of the course, he is also questioned about the challenges he faced, areas of strength and of improvement for each dimension of the proposals by Cabero *et al*., as well as how satisfied he is with his own performance. This questionnaire consists of open response reagents, so that the advisor can express his perception in writing.

III. Reviews of evaluation instruments

Eight meetings were held with the staff that had a direct relationship with the advisors, as well as the Unit Coordinator, the advisor on statistics and high school personnel in online mode. In the sessions the writing and order of the items was modified, and some were eliminated or added, resulting in the instruments indicated in the previous stage. Stages II and III were iterated until the final versions were obtained.

IV. Creation of digital forms

Because the unit does not officially have personnel in charge of information technologies, it is a programmer who makes modifications to the Virtual Classroom and their support is required to place the questionnaire aimed at students on the platform. For this, it is necessary to request it through a system of reports.

On the other hand, the supervisor was in charge of writing the items of the observation guide and the self-evaluation questionnaire as a form in a Word file, keeping the guide to record their observations, and sending the questionnaire by email to the advisors for them to record their responses.

V. Evaluation instruments' administration

The evaluation instruments were administered at the end of the course, allowing a period of five days to respond. This period ranged from October 11 to 16, 2019.

In order to encourage students to participate the evaluation was publicized on social media, as well as through the tutors, who communicated the importance of evaluating and indicated the

This work is under a Creative Commons Attribution 4.0 international license.

dates to carry it out. The invitations were made two weeks in advance, explaining the procedure and emphasizing that each student should evaluate each teacher with whom he took a subject.

The teachers to evaluate were 9, who taught (in total) 12 subjects. As each student assessed each teacher per subject, 160 responses were obtained from the 258 expected.

The undergrads' answers were recorded in a data concentrate in .xls format of the Excel program, that contained the evaluated teacher and the feedback for each ítem. The self-assessment questionnaire was sent by email to each advisor, obtaining 12 responses, that is, from the total population (the self-assessment was done per subject). On the other hand, based on their weekly observations, the advisory supervisor filled out the observation guide for each of the advisors.

VI. Analysis of results

The person who assisted on statistics analyzed the items of the questionnaire addressed to students through different procedures in the SPSS software, in order to identify if the instrument is reliable, if the reagents discriminate and if they are grouped into the proposed dimensions, presenting the results in an inform. These tetsing agents were coded numerically, and treated as if they were scale values, not ordinal. The findings presented in the aforementioned document are summarized below, and for reasons of space the complete figures with the results are not included:

Reagent discrimination test: this analysis was carried out in order to establish how much the reagents can separate the positive responses from the negative ones, applying the t-test for independent samples, comparing those items with a negative rating towards the advisors (with a score below quartile 25) against those in which it was positively scored (scores above quartile 75). The results indicate that all items discriminate.

Reliability of the instrument: An instrument is reliable if it can produce the same or similar results when applied to populations equal or similar to the original, that is, it is consistent in the results of its applications. To know the level of reliability of the questionnaire, Cronbach's alpha coefficient was calculated, and the result was 0.991. As noted by Hernández *et al*. (2014), the closer the coefficient obtained to the value 1 approaches, it indicates greater reliability, so the instrument has high reliability.

Grouping of items: To determine the dimensions in which the items of the instrument are grouped, a Varimax rotation analysis was performed. The outcome demonstrated that the items are grouped in only two dimensions, when originally it was designed so that the grouping was in three, being as follows:

Figure 1

This work is under a Creative Commons Attribution 4.0 international license.

Item grouping

Factor	Items grouped	Denomination
Factor 1	1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 14, 15, 16,17, 18,19, 20, 21, 22, 23, 24, 25	Pedagogical aspects
Factor 2	8, 13, 26, 27, 28, 29, 30	Technological aspects

Source: Martín, 2019

VII. Instrument adjustments

The results described in the previous stage led to the modification of the formula to calculate the teaching performance index (TPI). Therefore, the formula includes the two resulting areas in the piloting of the instrument aimed at students, in addition to the administrative area that is evaluated by the observation guide that the supervisor qualifies. The formula is as follows:

$$IDD=(.50*AP+.25*AT+.25*AA)$$

From the previous formula it is observed that half of the grade corresponds to the pedagogical area, where most of the items of the disciplinary area are already included, and the remaining half is divided equally between the technological area and the administrative area qualification. The instrument aimed at students as presented in Figure 1.

After the previous decision, the teaching performance index was calculated for each advisor. These indicators were also included in the pilot test report, so that the supervisor can count on them to prepare the reports for each advisor.

With this in mind, the reports prepared by the advisory supervisor contain the general information of the advisor and the subject in question, a figure with the scores obtained in each area and in the performance index, graphs with the percentages of the satisfaction levels expressed by the students and a summary of the comments added by the students in the open question, which, if required, are spellchecked, or those in which the answer does not contribute something to the advisor, are suppressed (for example, some choose to add "nothing" in the response space). Each advisor is responsible for comparing these results with what's stated in their self-assessment and has the opportunity to discuss the results at a later meeting with the assessors.

VIII. Improvement actions

This work is under a Creative Commons Attribution 4.0 international license.

Due to delays in the analysis and reporting of results, a meeting with the advisors has not yet been held, nevertheless, a session to evaluate the results and consider improvement actions is scheduled to take place in January 2020.

4. DISCUSSION

Performance evaluation, as Chiavenato (2009) and Castillo (2007) point out, is part of the human resources processes. In the unit case, through the diagnosis it was detected that some of these processes are the units' responsibility, such as the nomination of advisors to be and the evaluation of teaching performance, while others are the responsibility of the human resources department of the university. It's necessary to verify the alignment of these processes so that there are no discrepancies.

The purposes of performance evaluation are similar among those of any organization: of administrative and development value, according to Gómez-Mejía *et al*., and summative or formative, in the case of educational institutions, as Izar *et al*. (2016) affirm. In this paper's case, it is observed that the unit aspires to give formative use to the performance evaluation, which is reflected in the instruments' design stages and consecutive selected model, but what the unit does with the results obtained plays a crucial role in classifying the use in formative, summative or, worse, none. Until the conclusion of this article, the objectives, methods and model are aligned, contrary to what Gómez and Valdés (2019) found, which is the inconsistency between them. The unit is expected to maintain this congruence by making use of the results for the development and training of teachers in virtual environments.

The focus of this evaluation is partial, as Rubio (2003) refers, since it only addresses teaching performance. It is understandable, since the unit is new and the teachers, for the most part, have more experience in the face-to-face modality than at a distance, and as Fainholc (2004) points out, the fact that a teacher is good in front of a group does not ensure that their performance in distance mode be good. However, precisely because it is a new unit, it requires more complete evaluations, involving other areas, in order to achieve a significant and permanent improvement. Although teachers have a primary role in the service offered by the unit, its success does not lie entirely in their actions.

The issue of who the evaluators should be is also discussed, deciding in this case to evaluate from different perspectives, including those who are directly related to the advisor (students and supervisor), but also himself. Being an evaluation towards a service offered by people, and evaluated by people, objectivity becomes very difficult to maintain evaluators have their own

This work is under a Creative Commons Attribution 4.0 international license.

values, prejudices and even fears, which can influence the valuation they issue; including multiple points of view helps increase said objectivity. In addition, this multiple source evaluation follows the trend of teacher evaluation models in distance mode presented in the introduction, which encompass different perspectives, mainly from students, but also from coordinators, program managers, the evaluated and even the results (Llarena and Paparo, 2006; Villar, 2008; Duart and Martínez, 2001; Esquivias, *et al*. 2009; and Cabero *et al*. 2018). In this case, it was decided not to include the results in the evaluation because they are considered multifactorial and not the advisors' direct responsibility.

Regarding the instrument, the construction of the items based on an established procedure (the one proposed by Hernández *et al*., 2014) made possible to integrate each subdimension with observable indicators, the latter point being difficult to evaluate virtual advisors. There's knowledge and skills necessary for an advisor who, due to distance between teacher and student, becomes less obvious; for this reason, it was very useful to include different perspectives on the subject of evaluation, since what's not observable for students can be something evident for those who supervise the teaching, or for those who carry it out.

The items of the questionnaire aimed at students were presented to be answered by means of a Likert scale, which follows an ordinal logic. However, the pilot test was run in the SPSS program expressing each item as a scale variable, not a nominal type. Hernández *et al*. (2014) comment that this practice is common for social studies, however, this techniques applications' still controversial. Hernández *et al*. (2014) point out three characteristics that an instrument must have: reliability, validity and objectivity. Regarding the first aspect, the test showed that the instrument is reliable, that is, in future applications the results will be consistent. With respect to validity, as mentioned by the authors, it branches into that of: criteria, content and construct; no test was made on the validity of the criteria, as it did not have a reference that measures the desired dimensions; the validity of content and construct had unexpected results, as the test indicated the grouping of the items not in three but in two factors, this may be because the reagents in one of the dimensions were so little that it was unable for the test to identify them. In different circumstances, as already mentioned in previous paragraphs, objectivity was sought by including different assessments on the advisor, but this decreased because the formula designed to calculate teacher performance did not include the perception of the individuals under evaluation. Nor were the other two instruments tested.

These evaluation instruments, as explained throughout the document, were based on a theoretical model, which is specifically distance mode oriented. However, some aspects of said appraisal in higher level educational institutions were also taken into account to carry out the

This work is under a Creative Commons Attribution 4.0 international license.

preparation: in accordance to the recommendation of Pacheco *et al*. (2018), a good advisor's characteristics were first defined, this by reviewing documents and with attending to the opinion of the working group participants. This point was useful for this exercise, since during the review of other universities educational models that offer distance degrees it was found that the profile of those who teach the courses is very diverse and covers different functions and tasks, so they tend to vary from one institution to another, making it essential to establish the characteristics first, so that what's evaluated coincides with what's actually implemented.

Another proposal for the evaluation of face-to-face teacher performance in SSI that was taken into account was that of Mateo (2000) on the evaluation process phases: ideation, development, implementation and meta-evaluation, which are reflected in the intervention program proposed, with the exception of meta-evaluation, that's expected to be enforced by the unit prior to new administrations of the designed instruments.

As far as the selected model to support the instruments goes, the one Cabero *et al*. (2018) introduce, has certain strengths and weaknesses. Its strengths include the incorporation of different perspectives when evaluating, which, as mentioned in this section, allows reducing risks, increasing objectivity, along with an inclusive evaluation where everyone has a voice. Furthermore, the proposed dimensions are general and applicable to a wide variety of virtual teacher profiles; two of these dimensions are even mentioned by Pacheco et al. (2018): disciplinary and pedagogical knowledge, which apply to both face-to-face and virtual modality, adding in the latter the technological and administrative aspects, due to the fact that it develops in digital scenarios and it is essential to master tools of this field, in addition to the fact that these same scenarios make possible a closer follow-up to the teachers' work, since evidences are protected and it is possible to know the precise amount of participation in their virtual classroom, the activities they focus on, the interactions they have, etc.

Among the model's weaknesses are the shallowness of the proposal and the dimensions' focus on knowledge, setting aside personal aspects. The first point refers to the fact that the proposal is not extensive and is limited to stating the dimensions and subdimensions, leaving these factors to the interpretation of the reader, which in this case may not be completely correct because: a) the authors of the model are from a different country from where it was applied, and b) the wide diversity of distance teaching profiles, since it can be understood "at one's convenience". Secondly, the dimensions' focus on the teacher's knowledge can leave out internal aspects of himself that could be useful to provide a warm and friendly attention to the student (essential due to the lack of face to face contact), in addition to the attitude towards this modality (not very popular amongst teachers).

This work is under a Creative Commons Attribution 4.0 international license.

5. CONCLUSIONS

The three instruments resulting from this study, together, allow us a wider vision of teacher performance. The information is enriched and can be used advantageously for the evaluated teachers by including items that can be accounted, open responses, as well as coming from the perspective of the direct beneficiaries of the teaching process (the students), the one who continuously supervises the performance, and the evaluated himself. and the institution, which will obtain valuable information for decision making regarding improvement -decisions expected to be consulted with those involved.

Although this first exercise allowed the creation of instruments and their application, it is important to continue improving both of them during subsequent administrations, as well as opt for tools that reinforce their validity. In addition, the evaluation process needs to be constant and give a space of discussion to those involved.

The unit that was worked with displayed interest in evaluating teaching performance, in addition to the fact that the approach intended to give the results is the performance improvement of the virtual advisors, who have little or no experience in distance modality. These factors allowed and encouraged those involved to actively participate in the design and implementation of the instruments.

BIBLIOGRAPHICAL REFERENCES

Alcón, E. y Esteve, F. (2017). La importancia de la evaluación docente. *Cuadernos de Pedagogía* (476). 67-70. Recuperado de: http://repositori.uji.es/xmlui/bitstream/handle/10234/172538/52986.pdf?sequence=1&isAllowed=y

Cabero, J., Llorente, M. C., y Morales, J. A. (2018). Evaluación del desempeño docente en la formación virtual: ideas para la configuración de un modelo. *Revista Iberoamericana de Educación a Distancia 21* (1), 261-279. DOI: http://dx.doi.org/10.5944/ried.21.1.17206

Careaga, A. (2001). La evaluación como herramienta de transformación de la práctica docente. *Educere, 5*(15), 345-352. Recuperado de: https://www.redalyc.org/articulo.oa?id=356/35651519

Castillo, J. (2007). *Administración de personal. Un enfoque hacia la calidad.* Bogotá: Ecoe Ediciones.

Chiavenato, I. (2009). *Gestión del talento humano.* México: McGraw-Hill.

Chiang, M. y San Martín, N. (2015). Análisis de la Satisfacción y el Desempeño Laboral en los funcionarios de la Municipalidad de Talcahuano. *Ciencia & trabajo 17* (54). Recuperado de: https://scielo.conicyt.cl/pdf/cyt/v17n54/art01.pdf

This work is under a Creative Commons Attribution 4.0 international license.

Duart, J. y Martínez, M. J. (2001). *Evaluación de la calidad docente en entornos virtuales de aprendizaje.* Recuperado de: http://reddigital.cnice.mec.es/6/Documentos/docs/articulo12_material.pdf

Elizalde, L. y Reyes, R. (2008). Elementos clave para la evaluación del desempeño de los docentes. *Revista Electrónica de Investigación Educativa 10* (número especial) 1-13. Recuperado de: https://redie.uabc.mx/redie/article/view/198

Esquivias, M. T., Gasca, M. A. y Martínez, M. E. (2009). Competencias del docente virtual universitario y a distancia: sistema de evaluación por rúbricas. Trabajo presentado en *X Congreso Nacional de Investigación Educativa*, Veracruz, México. Recuperado de: http://www.comie.org.mx/congreso/memoriaelectronica/v10/pdf/area_tematica_02/ponencias/1657-F.pdf

Fainholc, B. (2004). La calidad en la educación a distancia continúa siendo un tema muy complejo. *Revista de Educación a Distancia* (12). Recuperado de: https://revistas.um.es/red/article/view/25311/24591

Fernández, E. y Luna, E. (2004). Evaluación de la docencia y el contexto disciplinario: La opinión de los profesores en el caso de ingeniería y tecnología. *Revista Mexicana de Investigación Educativa 9* (23), 891-911. Recuperado de: http://www.comie.org.mx/documentos/rmie/v09/n023/pdf/rmiev09n23scB02n04es.pdf

García, B. (2003). La evaluación de la docencia en el nivel universitario: implicaciones de las investigaciones acerca del pensamiento y la práctica docentes. *Revista de la Educación Superior 127* (32) 63-70. Recuperado de: http://publicaciones.anuies.mx/pdfs/revista/Revista127_S3A4ES.pdf

García, J. M. (2003). Los pros y contras del empleo de los cuestionarios para evaluar al docente. *Revista de la Educación Superior 127* (32) 79-87. Recuperado de: http://publicaciones.anuies.mx/pdfs/revista/Revista127_S3A6ES.pdf

Gómez-Mejía, L., Balkin, D. y Cardy, R. (2001). *Dirección y gestión de recursos humanos*. Madrid: Pearson.

Gómez, L. y Valdés, M. (2019). La evaluación del desempeño docente en la educación superior. *Propósitos y representaciones, 7* (2), 479 – 515. DOI: http://dx.doi.org/10.20511/pyr2019.v7n2.255

Guzmán, F. (2018). La Experiencia de la Evaluación Docente en México: Análisis Crítico de la Imposición del Servicio Profesional Docente. *Revista Iberoamericana de Evaluación Educativa 11* (1), 135-158. DOI: https://doi.org/10.15366/riee2018.11.1.008

Hernández, R., Fernández, C. y Baptista, P. (2014). *Metodología de la investigación.* México: McGraw-Hill.

Izar, J. M., Ynzunza, C. B. y Castillo, A. (2016). Propuesta de mejora al cuestionario de evaluación docente de la UASLP mediante análisis factorial y AHP. Trabajo presentado en *XX Congreso Internacional de Ciencias Administrativas (ACACIA)*. Mérida, México.

This work is under a Creative Commons Attribution 4.0 international license.

Llarena, M., y Paparo, M. (2006). Propuesta de una metodología de seguimiento y evaluación de cursos a distancia. *Revista Iberoamericana de Educación, 37*(4), 1-12. DOI: https://doi.org/https://doi.org/10.35362/rie3742702

Martín, M. (2019). *Informe de la prueba piloto del instrumento de evaluación del desempeño docente dirigido a estudiantes.* Datos duros inéditos.

Mateo, J. (2000). La evaluación del profesorado y la gestión de la calidad de la educación. Hacia un modelo comprensivo de evaluación sistemática de la docencia. *Revista de Investigación Educativa 18* (1), 7-34. Recuperado de: https://revistas.um.es/rie/article/view/121451/114141

Mishra, P., y Koehler, M. J. (2006). Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge. *Teachers College Record, 108*(6), 1017-1054. Recuperado de: http://one2oneheights.pbworks.com/f/MISHRA_PUNYA.pdf

Mondy, R. W. y Noe, R. M. (2005). *Administración de recursos humanos.* México: Pearson.

Mora, A. I. (2004). La evaluación educativa: concepto, período y modelos. *Actualidades investigativas en educación 4*(2). DOI: 10.15517/aie.v4i2.9084

Pacheco, M. L. C, Ibarra, I., Iñiguez, M. E., Lee, H. y Sánchez, C. V. (2018). La evaluación del desempeño docente en la educación superior. *Revista Digital Universitaria (RDU) 19* (6), 1-11. DOI:http://doi.org/10.22201/codeic.16076079e.2018.v19n6.a2

Páez, R. O. (2010). Evaluación de las funciones docentes en entornos instructivos virtuales (EIV). Revista Iberoamericana de Evaluación Educativa, 3(1), pp. 147-158. Recuperado de: https://revistas.uam.es/index.php/riee/article/view/4511/4935

Prieto, E. (2008). El papel del profesorado en la actualidad. Su función docente y social. *Foro de Educación* (10), 325-345. Recuperado de: https://dialnet.unirioja.es/servlet/articulo?codigo=2907073

Rico, M. I., Montalvo, E. Y. y Ayala, S. S. (2001). Utilidad de la evaluación del desempeño docente. *Revista de Enfermería del IMSS*, 9 (3), 137-141. Recuperado de: http://www.index-f.com/DOCUMENTOS_EBE/cuiden-evidencia/RE-IMSS-2001-3.pdf

Rosales, M. (2014). Proceso evaluativo: evaluación sumativa, evaluación formativa y Assesment su impacto en la educación actual. *Congreso iberoamericano de ciencia, tecnología, innovación y educación.* Buenos Aires, Argentina. Recuperado de: https://www.oei.es/historico/congreso2014/memoriactei/662.pdf

Rubio, M. J. (2003). Enfoques y modelos de evaluación del e-learning. Revista electrónica de investigación y evaluación educativa (RELIEVE) 9 (2). Recuperado de: https://www.fceia.unr.edu.ar/geii/maestria/2013/Dra.Roig/U3%20Rubio_ev_educacion_virtual.pdf

Rueda, M. (2018). Los retos de la evaluación docente en la universidad. *Publicaciones*, 48(1), pp. 143–159. doi:10.30827/publicaciones.v48i1.7334

This work is under a Creative Commons Attribution 4.0 international license.

Rueda, M., Elizalde, L. y Torquemada, A. D. (2003). La evaluación de la docencia en las universidades mexicanas. *Revista de la Educación Superior 127* (32) 71-77. Recuperado de: http://publicaciones.anuies.mx/pdfs/revista/Revista127_S3A5ES.pdf

Salinas, M. I. (2017). Gestión de la evaluación del desempeño docente en aulas virtuales de un proyecto de blended-learning. *Ciencia, docencia y tecnología 28* (54), 100-129. Recuperado de: http://pcient.uner.edu.ar/index.php/cdyt/article/download/222/275

Sastre, M. A. y Aguilar, E. M. (2003*). Dirección de recursos humanos, un enfoque estratégico.* Madrid: McGraw-Hill.

Vicario, C. M. (2015). El marco normativo de la Educación a Distancia: políticas y regulaciones. En Zubieta García, J. y Rama Vitale, C. (Eds.) *La educación a distancia en México: una nueva realidad universitaria,* 33-46. México: Virtual Educa. Recuperado de: https://www.ses.unam.mx/curso2016/pdf/18-nov-Zubieta_Educacion_distancia.pdf

Villar, G. (2008). *La evaluación de un curso virtual, Propuesta de un modelo.* Disponible en: http://www.oei.es/tic/villar.pdf.

Agradecimientos / Acknowledgments:
Expresamos el agradecimiento a nuestros docentes del Programa de licenciatura en educación de la Facultad de Educación de la Universidad Autónoma de Yucatán.

Conflicto de intereses / Competing interests:
Las autoras declaran que no incurre en conflictos de intereses.

Rol de los autores / Authors Roles:
Karla Ávila: conceptualización, curación de datos, análisis formal, adquisición de fondos, investigación, metodología, administración del proyecto, recursos, software, supervisión, validación, visualización, escritura - preparación del borrador original, escritura - revisar & amp; edición.

Leonor López: conceptualización, investigación, metodología, administración del proyecto, recursos, software, supervisión, validación, visualización, escritura - preparación del borrador original, escritura - revisar & amp; edición.

Fuentes de financiamiento / Funding:
Las autoras declaran que no recibió un fondo específico para esta investigación.

Aspectos éticos / legales; Ethics / legals:
Las autoras declaran no haber incurrido en aspectos antiéticos, ni haber omitido aspectos legales en la realización de la investigación.

This work is under a Creative Commons Attribution 4.0 international license.