
University teacher evaluation and professional development: a review based on participants, dimensions and methods

Evaluación docente y desarrollo profesional universitario: Una revisión basada en los participantes, las dimensiones y los métodos

教师评估与大学专业发展: 基于参与者, 维度和方法的回顾

Оценка учителей и университетское профессиональное развитие: обзор по участникам, измерениям и методам

Noemí Suárez Monzón

Indo-American Technological University of Ecuador (Ecuador)

nsuarez@unibe.edu.ec

<https://orcid.org/0000-0002-9103-9714>

Maritza Librada Cáceres Mesa

Autonomous University of the State of Hidalgo in Mexico (Mexico)

mcaceres_mesa@yahoo.com

<https://orcid.org/0000-0001-6220-0743>

Vanessa Gomez Suarez

Indo-American Technological University of Ecuador (Ecuador)

vgsuarez1@gmail.com

<https://orcid.org/0000-0001-8733-4190>

Isabel Cristina Pérez Cruz

Quevedo State Technical University in Ecuador (Ecuador)

iperez@uteq.edu.ec

<https://orcid.org/0000-0002-5091-9838>

Dates · Fechas

Received: 2021/09/01

Accepted: 2021/09/29

Published: 2022/01/10

How to Cite this Paper · Cómo citar este trabajo

Suárez, N., Cáceres, M. L., Gómez, V., & Pérez I. C. (2022). University teacher evaluation and professional development: a review based on participants, dimensions and methods. *Publicaciones*, 52(3), 161–185. <https://doi.org/10.30827/publicaciones.v52i3.22271>

Abstract

Teacher evaluation is a cycle that reveals forms of professional development. It is a participatory and multi-dimensional process. Therefore, the aim of this study was to analyse the different ways in which the participants' evaluations contributed to the professional development of teachers in different university settings. For this purpose, a literature review was carried in 6 stages that allowed us to analyse and identify the main findings of 48 studies in the international context. The topics covered in the literature were analysed and synthesized, and the following findings were obtained: 1) stressors were identified in relation to differences and imbalances in the contributions of each participant to teacher professional development; 2) all the participants evaluated predominantly similar dimensions essentially related to teaching and research, with little relevance attached to a significant number of teaching dimensions, characteristics and skills that contribute to teacher professional development; 3) the methods used were based on segmentation by participant without achieving methodological coherence to ensure a relevant evaluation of aspects of teacher training and professional development.

Keywords: teacher, students, evaluation, managers, peers, participants.

Resumen

La evaluación del profesorado constituye un ciclo que evidencia las formas de crecimiento profesional. Este es un proceso de naturaleza participativa y multidimensional. Por tanto, el objetivo es analizar las diferentes formas en que la evaluación de los participantes aporta al desarrollo profesional de los docentes en diferentes escenarios universitarios. Para lograr esto, se realizó una revisión de la literatura siguiendo 6 etapas que permitieron analizar 48 contribuciones en el contexto internacional y descubrir los principales hallazgos. Mediante una síntesis que recoge el análisis de contenido temático de la literatura se exponen los siguientes resultados: 1) existen tensiones en cuanto a la desintegración y desequilibrio en los aportes de cada participante al desarrollo profesional docente; 2) hay predominio en la evaluación de dimensiones similares por todos los participantes relacionados esencialmente con la docencia y la investigación, quedando sin relevancia un número importante de dimensiones, características y aptitudes del docente que aportan al desarrollo profesional; 3) existe un empleo de métodos que responden a la segmentación por participante sin que se pueda lograr una coherencia metodológica para lograr una evaluación relevante en aspectos de formación y desarrollo profesional del docente.

Palabras clave: docente, estudiantes, evaluación, gestores, pares, participantes.

摘要

对教职员工的评估构成了表明专业发展不同形式的一个循环。这是一个参与性和多维性质的过程。因此，本文目的是分析参与者的评估如何以不同方式对不同大学环境中教师专业发展做出贡献。为实现这一目标，我们通过以下六个阶段进行了文献回顾，在该过程中我们分析了国际背景下的48项贡献并获得主要的发现。我们通过对文献主题内容分析的综述得出以下结果：1) 在参与者对教师专业发展的贡献的分裂和不平衡上存在紧张关系；2) 所有参与者在与教学和研究相关的维度的评估中占主导地位。维度数量、做出职业发展贡献的教师特征和才能之间不相关。3) 存在一种方法可以对参与者的细分，但无法达到方法上的连贯性，从而无法在教师培训和专业发展方面进行相关评估。

关键词: 教师, 学生, 评估, 管理人员, 成对, 参加者。

Introduction

Research on teacher evaluation in higher education can be divided according to subject involved (academic peers, students, or teachers themselves) as the only participant in the process. Depending on the participants selected, in order to carry out the evaluation, the role of each one in the dimensions they are capable of evaluating and the methods used to collect this information would be determined. The definition of the dimensions to be evaluated and the methods to be used by each participant constitutes a stressor; the defined parameters are influenced by the educational macro-policies and characteristics of the universities. However, other studies have highlighted the value of participatory teacher evaluation in which all the participants play a prominent role. This would condition equilibrium in the contributions made by each individual to the professional development of university teachers. In this way, a comprehensive overview of the role of teachers would be obtained.

With respect to the so-called fractional participation approach, involving only one participant, students play a leading role in the teacher evaluation process. Thus, evaluation is conceived as a supervision, oversight and job promotion strategy used as a retaliatory mechanism against teachers (Hornstein & Edmond, 2017). Anonymous questionnaires are often used as a method for gathering information. This affects teacher performance and their motivation to develop professionally. They have a low impact on the actual feedback provided for their work to the extent of disrupting and modifying the teacher's behaviour (González et al., 2016).

Another approach, adopted by some authors, includes teacher self-assessment, i.e. the teachers themselves establish critical criteria on their behaviour and contribute to their personal-professional development, thus fostering greater intrinsic motivation. Self-assessment or self-evaluation helps teachers to find new forms of self-development to meet the needs of their students and fulfil the university's mandate in social development as a criterion of quality accreditation policies. In this context, the direct participation of teachers in their own evaluation, i.e. the entire teaching dimension, mainly through questionnaires, is a key aspect as shown in the study by Ruiz-Corbella and Aguilar-Feijoo (2017).

Other participants who are also prioritised are academic peers. Currently, the critical and reflective approach to teaching practices is described in many studies due to its relevance for teacher improvement. Some evaluation practices are normally carried out through observations with peer teachers. They generally have common communication codes as they belong to the same or related areas of knowledge, and they are able to technically assess the specificities of the teaching-learning process in a particular discipline or field of knowledge (Contreras, 2018; Olive & Costa-Lobo, 2019).

In contrast, less attention has been given in research to the participatory approach involving several individuals. The studies describe different interactions (teacher-peer; student-peer; teacher-peer-student). With respect to our study, due to the participatory nature of the evaluation, it should be treated as a research process guided by managers, which also adds high levels of conceptual and methodological complexity due to the certainties that must exist between the evaluator, what is to be evaluated and how the evaluation should be performed. This implies establishing a series of dimensions, methods and subjects for the professional development and improvement of teachers (Escudero, 2019).

Despite the large number of studies on teacher evaluation in the literature, we did not find any review of the literature describing the state of knowledge from the perspective of the participants. Therefore, the following research question was posed in this paper: What contribution do participants make to the professional development of university teachers?

Therefore, the aim of this study was to analyse the different ways in which the participants' evaluations contributed to the professional development of teachers in different university settings.

Method

After defining the objective of this study, a literature review was carried out using the stages proposed by Ferreira et al. (2011) adapted for this study: 1) definition of the question of interest; 2) identification and selection of relevant studies; 3) extraction of data from primary studies; 4) selection based on the study inclusion and exclusion criteria; 5) analysis of the selected texts; and; 6) presentation of the results.

This was carried out by researchers from two universities who share similar concerns about the evaluation of university teachers, in three stages: in the first stage, the texts were selected; in the second, the roles of researchers were distributed to process the data from the selected texts; and in the third, the selected texts were analysed to determine the contribution of the participants in terms of the dimensions evaluated by them and the methodologies used for teacher professional development.

To define the population of articles to be considered, only the main researcher carried out the initial search to avoid duplication. Only articles with titles containing the keywords "evaluation", "self-evaluation", "peers", "teacher", "university", "higher education" or "training" in English and also the Spanish equivalents in Google Scholar and Eric were selected. The Boolean connectors AND and OR were used in different combinations of the aforementioned terms. A total population of 160 articles was established, comprising articles in Spanish and English from peer-reviewed journals.

The main researcher consolidated the information in a matrix, which was completed with the year of publication, name of the journal, quartile, title of the article, country where the research was carried out, and the abstract.

Then, each researcher analysed the content of the abstracts of 40 articles to determine the final sample. Their acceptance was defined according to the most significant inclusion criteria according to the objectives of the study, which are detailed below: 1) the article had to specify the person evaluating teacher performance (student, academic peer, teacher, manager or various subjects); this criterion was used due to the importance of the participatory-collaborative approach for the evaluation (Jara & Díaz López, 2017); 2) the article had to contain the following: dimensions/evaluation criteria, technical methods or evaluation instruments. These aspects were used due to current concern regarding the procedures and instruments used in evaluations (Navarro & Ramírez, 2018).

At the same time, the following exclusion criteria were taken into account: 1) evaluation, specifically the one used by teachers to evaluate student learning; 2) evaluation concepts (since the analysis of the evaluation category was not an aim of the study); 3) participants not considered key to this research (other than students, peers, teachers,

managers); 4) articles that, despite focusing on teacher evaluation, were carried out in non-university contexts because our analysis focused on higher education teachers.

The review yielded a total of 48 articles published in scientific journals in the period between 2008 and 2020, which proposed peer reviews for the publication of articles, indexed in different quartiles, according to the data obtained from the Scimago Journal & Country Rank portal, which confers a high level of reliability. Most of the articles in this sample were in English. The studies had been carried out in various countries, with Europe being the most represented continent. The participants in most of the studies were mainly students from multiple university systems, with a greater presence of articles describing peer evaluations. Self-evaluation and evaluations carried out by several people were less represented.

Then, each researcher analysed the specific thematic content in the context in which they were distributed by the main researcher, which meant that each researcher analysed 12 documents (Fernández, 2002). Subsequently, a narrative synthesis of the findings was carried out (Davies et al., 2014) to inform the dimensions and/or methods to be taken into consideration by one or more participants in the evaluation, supported by the frequency table.

This section summarises the findings based on four main axes, which were identified with the participants in the evaluation: students, teachers, peers, and various participants. These axes were used to classify the dimensions evaluated by each participant and the methodologies used to present their opinions. This association was made according to the teachers' ability to perform professional functions in higher education (teaching, research, management) (Ruiz et al., 2008), as well as their characteristics and attitudes (Suárez et al., 2018). The methods or techniques used in each case were also analysed. The narrative synthesis presented in this paper provides details of the contributions of each evaluated dimension and the methods used to contribute to teacher professional development.

Results

This axis shows the teacher evaluations performed by the students, who indistinctly evaluated at least four teaching-related dimensions: teaching methods, communication, research and personal-attitudinal characteristics. The contributions clearly revealed purposes that favour the professional development of teachers to a greater extent, but also less research on teaching and the characteristics and skills of teachers.

In this review, Spain was the main contributor of articles focusing on evaluation with students as participants, deriving from the "Docentia" programme proposed by the National Agency for Quality Assessment and Accreditation ("Agencia Nacional de Evaluación y Acreditación de Calidad" - "ANECA" - in Spanish). A student questionnaire is normally used comprising three dimensions - planning, development and results - that could fall within the scope of the management of the teaching-learning process (Moreno-Murcia et al., 2011), due to its key role in the maintenance of university quality (Hortigüela et al., 2017; Jornet et al., 2011). From our detailed analysis, there seems to be great interest in methodologies, the use of novel resources and evaluation systems to achieve learning and their application in future professional contexts, which are all relevant aspects in the progress of students but not necessarily in that of teachers (Bilbao & Villa, 2018; Dios et al., 2018; Juanas & Beltrán, 2014; Eizagirre et al., 2017).

However, certain autonomous communities (regions) in the country re-contextualised this model and introduced modifications. One such modification was implemented at the University of the Basque Country, which adapted this questionnaire and included the following dimensions: relationship with students, innovation and improvement, and overall satisfaction with a series of items ranging between 18 and 28 (Lizasoain-Hernández et al., 2017). The intention was clear, namely for students to evaluate methodological innovation. This point is central to the improvement of the learning process of both students and teachers and sets a precedent for the teacher training necessary to engage in critical reflection and improvement exercises through pedagogical research.

In countries like Belgium, Japan and Canada, students' opinions are obtained through questionnaires and are used mainly to evaluate the use of technology in teaching and, above all, to make administrative decisions in teacher promotion. This approach has been strongly criticised (Spooren et al., 2013; Spooren & Van, 2012).

Technological competences are clearly necessary in current times in which distance learning has become universal in university education. However, very often students, enthusiastic about computers used by their teachers, neglect to assess the effectiveness of their learning. The pedagogical use of computers must therefore be evaluated. It is also not entirely clear to what extent students are able to evaluate all aspects of teacher competence to confirm that this is the criterion that determines teacher promotion (Hornstein & Edmond, 2017).

In Mexico, Colombia, Chile and Ecuador, evaluations are also conducted using questionnaires with great emphasis on supervising teaching competence in terms of planning, evaluation and communication skills in order to achieve student satisfaction and to manage and assess the teaching-learning process. All this is the result of the application of traditional teaching methods and is associated with the dimension of methodological innovation, which is more and more present in recent literature (Donado et al., 2018; Escobar, 2015; González et al., 2016; Luna & Reyes, 2015; Márquez & Madoño, 2016; Serrano et al., 2019). In the Latin American context, there is a similarity with Spain insofar as students evaluate teaching as a whole for supervision purposes to improve student learning, as well as the methodological innovation associated with teacher training processes.

It is known that the approach to teacher training in both Spain and some Latin American countries is aligned with the results obtained in these dimensions by many different centres and institutions (Mena et al., 2019). However, it is uncertain whether student assessment of methodological innovation actually fosters a restructuring of the teacher's field of action and consequently leads to research and subsequent publications on the results obtained, which would indicate quantifiable progress on teacher reflection and innovation in teaching.

In fact, student assessments of the abilities to work with scientific information and develop projects with students are mentioned in the literature, albeit with few articles focusing on this subject. This result seems to be associated with the prioritisation of teaching according to the strict definition of classroom-based teaching and not in its relationship with research because the evaluation of these aspects is not as important in some regions. As in Trinidad and Tobago and Ghana, the classroom components are still the main dimensions evaluated through questionnaires, together with teacher's communication skills and personal characteristics (Blair & Valdez, 2014; Nyame et al., 2019).

In contrast to the aspects analysed so far, the study by Simon et al. (2017) in Slovenia reveals that teachers are evaluated using a questionnaire that measures eight dimensions, all related to teaching quality, but its relevance is marginal with respect to research measured by counting the number of impact articles published by teachers. The results of this study showed that publication quality is related to teaching quality. While Jalbert (2019), in a sample of 300 business faculties at 104 universities in the United States, did not find evidence to indicate that research quality influences teaching quality, these two dimensions seem to be unrelated, at least according to the scores on the scales applied to students.

According to Hortigüela et al. (2017) some students believe that teachers' dedication to research reduces their time to prepare innovative classes. Teachers are recognised more for their methodological performance in the classroom and their availability than for their prestige or rank. However, teachers have the opposite opinion; they believe that the category and academic and research prestige of the teacher are real indicators of higher teaching quality. These results highlight the need for teachers to conduct pedagogical-didactic research to innovate in teaching-learning processes in the corresponding field of knowledge. As the author states, both teaching and research and innovation are important for the personal training processes of both students and teachers. Thus, attributing a balanced weight to the link between teaching and research is extremely important.

The presence in the articles studied of the dimension of affective communication between teachers and students, as well as the development of creative and innovative thinking, is relevant. Both reflect basic principles of the teaching-learning process, such as coexistence, respect, trust, and warm and affable relationships that make teachers more "accessible" as demanded by students to facilitate the acquisition of knowledge. Student-teacher interaction is special when there is communication between people with different interests, ages and levels of reasoning (Escobar, 2015; Pomp & Pérez, 2015). Therefore, in these terms, it is reasonable to think that the evaluation of this dimension by students would serve as an intrinsic motivation for teachers to prepare themselves for the use of creative-recreational strategies that satisfy the learning needs of students and enable them to assert their positions as teachers. In contrast, the use of authoritarian communication mechanisms based on the power relationships of teachers over students can influence unfavourable evaluations; hence, their contribution to teacher development disappears.

Another equally important but under-represented dimension, in the opinion of the researchers in the sampled studies, is the evaluation of teachers' characteristics and attitudes. According to Escudero (2019), although this does not absolutely guarantee quality in education, it is a necessary condition that can be evaluated by students. Whatever the case may be, this aspect received little attention in studies published in different countries, and some characteristics or qualities were scarcely evaluated by students, including the following: enthusiastic, fair, respectful, creative, humble, flexible, up-to-date, collaborative, responsible, autonomous and dedicated.

In general, regardless of the different objectives of each study, it was evident that almost all of the studies emphasize the need to evaluate the different aspects of teaching and, consequently, teaching competences from planning to results obtained by students, as well as affective communication and development of critical thinking to demonstrate the successful practices of teachers. The emphasis is placed on student results, which reveals the emphasis on the objectives of control, performance oversight and, in most cases, teacher promotion (Eizagirre et al., 2017). Although the inno-

vation dimension was present and implies research, the possession of skills to work with information and develop projects with students is not sufficiently evident in terms of the development of the research abilities of both students and teachers, since the aspect evaluated in this case is impact research. The stressor corresponding to the low presence of this dimension in teacher evaluation increases concern when considering its contribution to innovation, the educational leadership skills of the teacher and the development of generic or professional development skills (Suárez et al., 2018).

This section of our research addresses the self-assessment of university performance and teachers as active participants. It includes the dimensions associated with self-reflection on teaching, innovation, communication and management as aspects that promote professional teaching development and university quality.

The study carried out by Loredó et al. (2008) in Mexico declared that research on performance evaluation, carried out by teachers themselves, was either scarce or not recurrent in literature. Hence, the aforementioned authors conducted a study consisting of interviews to gather the opinions of the evaluated teachers and based on the theoretical characteristics of a good teacher to create the bases of an evaluation model. These authors insisted that it must precede all other forms of evaluation of the effectiveness of teaching to tailor teaching to the learning of all students. This is connected with the idea that teachers are the best judges of their performance provided this is evaluated in a responsible and mature manner.

Based on this assumption it follows that criticism and reflection regarding the approach were present in all the articles studied. It is agreed that professional and personal progress is achieved when teachers are required to self-reflect on performance, but we believe that it should not necessarily precede other forms of evaluation; on the contrary, it must happen. This means that teachers are able to comprehend their own performance more maturely when they reconstruct their practices and receive inputs from students and peer evaluators.

Researchers at the University of Tartu in Estonia followed this approach, using the reflective practice questionnaire to recognise the role of reflection in teaching, learning and personal experiences in professional development (Kalk et al., 2014). Likewise, the use of the questionnaires with English teachers from 4 universities in Iran revealed the teachers' feelings, experiences and practices regarding teachers' self-reflection on classroom management (Sammaknejad & Marzban, 2016). In fact, teacher self-reflection on their own activity will allow them to improve performance because it helps them identify successes and failures in their professional activity in terms of their theoretical position in the discipline, challenges and their pursued goals.

From this perspective, reflection changes the focus from the personal level to the group level and requires teachers to relate with their peers, thus driving professional development. However, the authors showed that in order to evolve efficiently and improve their performance index rating, it is necessary to implement programmes similar to those in countries such as Finland and Singapore, which integrate other methods such as the portfolio method and the observation of videos to de-construct teaching actions in the classroom and teacher diaries on specific events. For example, evaluations by students, relationships between teachers, behaviour control or family-school relationships. These aspects become very important for feedback and professional growth.

In Russia, Tyunnikov (2016) provides guidelines for this process of self-reflection to foster the transformation of the teaching process through innovation, using a ques-

tionnaire with items relating to motivation for innovation, teaching planning, experimental management and communication of results. An interesting aspect of this study is that it arouses interest in research and allows teachers to gradually develop investigative skills in the innovation process autonomously. In Ecuador, Ruiz-Corbella and Aguilar-Feijoo (2017) developed a questionnaire to evaluate the comprehensiveness of teaching, research and management based on the self-reflection of strengths and weaknesses to determine the specific needs to be addressed in teacher training programmes.

The aforementioned analysis revealed the strong emphasis placed on critical thinking and reflection in the contributions, which may be due to the demands of pedagogical innovation as a key element to improve teaching and as a consequence of the development of teaching, communication and pedagogical management skills. This study presents research associated with the innovation of the teaching process. The self-evaluation of personal characteristics and attitudes was not observed, which is understandable to some extent due to the difficulty involved in their self-recognition, since it is often a cause of tension as it is not expected by students and less so by managers.

Peer evaluation can be conflictive because some studies show that peers do not contribute to professional development, but most authors maintain that peer evaluation is of special relevance for the improvement of teaching. They allude to the fact that peers can better judge teaching practices, the relevance of the use of teaching methods or actions for a certain discipline or field of knowledge in particular. Therefore, it can be used to stimulate the professional development and improvement of teachers and consequently students.

Levander and Riis (2016) in Sweden were unable to demonstrate that peer review had a positive impact. They evaluated the educational dimension that included teaching experience, reflection on teaching topics and research differentiated by the type of publication. These authors studied the records and presentation of educational philosophy of 294 Swedish teachers promoted in different disciplinary domains. They reported differences in how they described and judged teachers in different disciplines and that teacher evaluation did not fulfil its formative function, nor did it generate quality. Despite the fact that the peers had the same academic background, they should have known how to carry out the evaluation process to avoid undermining formative influence.

Along these lines, Yiend et al. (2014) carried out a study in the United Kingdom using a hybrid teaching observation model, in response to doubts expressed in previous studies on the preparation of peers to perform evaluations, the contribution of feedback and to determine whether evaluation is really able to catalyse teacher professional development. This is a professional orientation and development model that incorporates an expert-researcher and a collaborative teacher who frequently monitor each other in the context of the discipline.

The results showed that teachers modify their teaching practice according to the motivations and feedback provided by students. However, in the authors' opinion, designing peer evaluation programmes structured around this model has negative consequences because the pressure of an experimental evaluation conducted by an expert can give rise to stressful situations, but also positive situations because it is a formative path for teachers from disciplines other than Pedagogy, while at the same

time allowing teachers to become involved in research by interacting with an expert in their area.

A systematic review by researchers from Malaysia and the United States Thomas et al. (2014) revealed the barriers and benefits of peer reviews in the development of faculty members, but the descriptive synthesis revealed the need for such teaching evaluations to include other aspects because they can hinder the evaluation process, minimize the fear of being seen by others in the same area, the negative impact of feedback because criticism is still seen as being detrimental to professional development, distrust in the experience of the evaluator, as well as clarify the understanding that evaluated individuals and peer evaluators have of effective teaching.

However, despite the barriers, these authors highlighted the benefits associated with what they considered to be one of the most beneficial dimensions: self-reflection and reflection on peer feedback because it helps teachers from different disciplines to understand, relate, learn and impart teaching philosophies and in an amalgam of inter-cultural collaboration in different academic contexts to innovate teaching practice.

A sui generis process and tool for performing peer evaluations in a medical school in the United States was presented in the study by Wellein et al. (2009) to evaluate teaching (objectives, organization, content, presentation style and relationship with students) and document teacher performance. In this case, the evaluated teachers and evaluators were prepared beforehand using a document outlining the teaching characteristics sought by the most experienced teachers. The Likert Scale used by three different participants was approved by all the members of the faculty. This scale assigned a key role to the participation of all and their holistic vision of teaching performance was recognized from the beginning to the end of the process.

Although the authors and participants recognised the value of the process, they also acknowledged the need to complete the information from different perspectives, which means that not only peers but also teachers should evaluate and the academic leaders of each group should organise the scientific discussion around the discipline and the development of their teaching. The relevance of this idea lies in taking into account the particularities of areas of knowledge even when there are general characteristics regarding what is expected of teaching in higher education. Establishing a single evaluation system by areas of knowledge would be particularly relevant in teacher professional development.

Following this line of collaboration, in Chile there is a strong demand to construct didactic knowledge through dialogue between peers, although this parity does not occur in the broad sense and the objective is dissipated due to the hegemony of one group over another. For this reason, it is normally organized to foster a discussion that articulates academic identities and professional knowledge to address the complexity of consensual professional practice eventually resulting in growth and improvement. Recently, formative peer assessment was considered a valid approach when an experienced and informed teacher guides professional development through significant and more complex elements relating to management and feedback that improves teaching (Contreras, 2018). This author proposed a work cycle between peers sharing the same areas of knowledge comprising the following: 1) a pre-observation meeting; 2) a record of the observation of teaching methods in class; 3) a report; 4) a post-observation meeting; 5) a final report centred around the experience of encouraging teachers to reflect on their teaching and the learning results achieved by their students.

In Canada, a more recent review by Hamel and Viau-Guay (2019) included the use of video to support the development of reflective practice in initial and on-going teacher training. They identified significant aspects in the role of these interactions in the practices of in-service teachers who engaged in collaborative reflection with peers. The studies indicated that teachers engage in a more solid reflection because they focus more on students and their learning than on themselves. At the same time, the value of collective teacher learning fostered in the teaching community in turn fosters the critical awareness of the participants to understand the degree of development of their abilities. Evaluating both individual and collective reflections on teaching through the observation of videos, as a training and development mechanism for teachers, represents a change in approach compared to traditional course-based training. However, this puts into perspective the level of teachers' interpersonal relationships with their peers to freely define and document this learning from the point of view of the actors themselves.

Three studies conducted in Australian universities examined the peer review approach given its special connotations for teacher professional development and teaching excellence. Peers were defined as "academic colleagues who give and receive feedback on teaching" (Grainger et al., 2016, p. 1), thus leaving open the possibility of dialogue on equal terms in a model of peer-assisted teaching. Georgiou et al. (2018) provided guidelines on pedagogical and didactic aspects of learning and teaching. A distinctive element in these studies is the sequential approach to integrating different methods, such as the workshop, observation, meetings and the development of an improvement plan based on the practice of reflective-criticism and collaboration in the mentoring of colleagues. This aspect has an impact on prospective development because the planning of improvement actions puts the teacher in a situation of change, especially when the areas of improvement are understood and internalised.

The narrative analysis of 20 university professors in Portugal also supported the work of interaction between peers due to its important contribution in the process of feedback to teachers. A special form of improvement is achieved through the affective experience of treating each other as equals, exposing certainties, fears, thus marking the difference between the routine control of teaching and supervision (Olivera & Costa-Lobo, 2019). This information is recorded by the manager responsible for the purpose of on-going improvement and the contribution of peers to reflection, feedback and consequently improvement of student learning in the classroom.

It follows from this analysis that, despite the evident discrepancies between the different authors and the necessary conditions for organisation of peer evaluation, the findings tip the scale in favour of the use of peer evaluation as one of the most valuable methods for transforming teaching practices and promoting teacher professional development.

The higher-level evaluation by several of the participants mentioned previously enables better teacher professional development.

Elizalde and Reyes (2008) offer their participatory vision of evaluation based on the creation of programmes that gather data compiled from various sources: students, academic peers, and teachers themselves. In their opinion, peer evaluation should be carried out on stressful aspects such as reflection on learning context and the factors that promote it, the collaborative work processes and systems involved, the discussion of proposals, alternatives to implement, incorporate and evaluate changes in practice and produce resources necessary for student learning. The emphasis was placed on

the nature of the process agreed between the teacher-peer and the conditions that the manager must prepare to provide spaces for pedagogical discussion.

In these conditions, it is necessary to create an immediate environment of motivation among colleagues who can professionally discuss pedagogy and didactics. This favours the exchange of experiences, good teaching practices and the philosophy of each individual. In the context of university professors, this communication exchange can give rise to conflictive situations, agreements and disagreements, but ultimately a contradiction that fuels the development of the participants.

Xu (2012) in China, Mas (2014) in Spain and de Jara and Díaz López (2017) in Colombia followed the same approach with respect to the use of multiple actors to complement student evaluations. According to Xu (2012), the development of a feedback system using multiple sources of data from different actors could be critical but at the same time constructive. Documenting performance and providing a comprehensive view of teacher performance requires student ratings, peer ratings, teacher self-assessment, administrator ratings, teaching awards, and more. The tension identified in their study stemmed from the way of establishing a multi-dimensional evaluation system integrating the merits of different evaluation approaches with reliable, effective and equitable data in the evaluation system.

The content of this study refers to the dimensions of teaching related to the methodologies used by teachers that should be evaluated by students and introduces the idea that peer evaluation should be performed by other institutions, which undoubtedly reduces the bias produced by the proximity of teachers in the same faculty, but limits the communicative and empathic relationships necessary to engage in academic, scientific and pedagogical discussions. In the authors' opinion, evaluation processes performed by specialists outside the university can be stressful and not comply with the spirit of teacher discussion and transformation.

The inclusion of Spanish universities in the European Higher Education Area (EHEA) modified the assigned functions and thus the requirement to acquire and/or develop new competences to respond to these demands. Research was one of the competences constructed by Mas (2014) who defined the competence profile of teachers: design, develop and evaluate research projects; organise and manage scientific events; develop scientific material; communicate and disseminate scientific advances. For this purpose, the author conducted a study in which he triangulated different sources of information (teachers, experts and students). This investigative practice on the part of the author, based on the construction of the profile by the participants in the evaluation, empowers the latter to transform performance. It implies criticism, respect and, above all, group participation as a resource for individual development.

In the case of Colombia, according to Jara and Díaz-López (2017) the requirement to fulfil certain functions in the gathering of evidence to measure the results of student, teacher and institutional performance adapted to the 21st century should, in many countries in the region, give way to critical and reflective rationality of university processes based on participatory evaluation. This concept is encompassed within the context of the policies implemented by each university to evaluate performance based on multiple functions where all have the same importance. This idea of equating the corresponding weights reveals the systemic nature of substantive functions, attributing more weight to one than to others necessarily makes for an uneven process that negatively influences teachers, students and the university as an institution. Another relevant idea in this study is that the participants, together with the teacher, generate

pedagogical knowledge and permanent feedback, allowing the teacher to learn and grow with them.

The experience of the University of Laguna in Spain in teacher evaluation also supports the participatory approach. The *Docentia-ULL* programme is carried out through the triangulation of three sources of information: a student satisfaction questionnaire; reports from their academic managers; and teacher self-reporting on his/her own merits in the three dimensions of the model proposed by ANECA at national level (teaching planning, teaching development and results and innovation). The model enables a better distinction to be made between excellent, noteworthy, compliant, and non-compliant teachers. There was an imbalance in the priority attributed to different dimensions within the teaching function. Based on the modifications made to the weights assigned to each dimension of the teaching function, it was possible to assess the evaluation of teachers' commitment to training and educational innovation and reduce the weight assigned to the evaluation of bureaucratic teaching obligations, while the results of the satisfaction survey conducted with students and academic managers has become a requirement independent of the merits of the teacher in the evaluation (Isla-Díaz et al., 2018).

Indeed, this study follows the approach adopted in the study described previously with respect to the weight assigned to the different dimensions. Replacing the assessment of teachers' commitment to training and educational innovation necessarily means that more attention must be paid to pedagogical research and innovation in the classroom, which could modify the unfavourable opinion of students reported by Hortigüela et al. (2017) with regard to the time that teachers devote to research and teaching quality. Another significant aspect is the allusion to the participation of academic leaders or managers in teacher evaluation, which, in the context of this review, was only mentioned in the studies by Wellein et al. (2009) and Xu (2012); this aspect seems vital not only in terms of the participation of these individuals as evaluators of certain dimensions but also for bringing together all the participants in a coherent manner.

Finally, Escudero (2019) describes the diversity of areas of action, contents, and objectives of the evaluation employed in the different approaches, and acknowledges that it is a particularly complex task conceptually and empirically speaking. In view of the foregoing, the aforementioned author proposes a series of methods, sources and strategies to be adopted by the different actors in the education community. This study confirmed the participatory, investigative and formative nature of teacher evaluation as a premise for professional and educational improvement.

This analysis also reveals the indistinct contributions between teacher-peer-students, all as key subjects in the teacher evaluation process. Certain studies clearly reveal the dimensions they are evaluating, while others are less clear because they refer more to the actors who are required to complete the data and the methods necessary for that purpose. However, a necessary organisational condition from a practical standpoint is that managers must be able to scientifically guide this process; the manager's unifying role is not clear, hence the need for an investigative roadmap to guide, both in theory and in practice, the coordinated participation of all subjects. The managers have the capacity to bring the participants together, collect data and triangulate information to identify teacher training needs and, at the same time, their potential for the benefit of the group.

The literature studied clearly shows the segmentation and bias in data collection, obtained from the different persons evaluating teachers, and also the tendency to make this process more complex due to the wide range of quantitative and qualitative methods. In turn, there is limited evidence of the data triangulation process in the participatory evaluation phase.

In the teacher evaluations by students analysed in the articles studied, the quantitative questionnaire was invariably the instrument used in the teacher evaluation. For some authors this is highly relevant due to its key role in maintaining university quality (Jornet et al., 2011). A study in Portugal confirmed the reliability and consistency of this method (Moreira & Santos, 2016). The authors demonstrated that the use of the *Student's Evaluation of Educational Quality (SEEQ)* tool was suitable for the evaluation of educational quality by students. This was validated by students and teachers, revealing the consensus between them.

In contrast to these findings, many studies have shown the reasons for its low relevance in the improvement of student learning and its low level of use for professional development purposes. One of the many reasons is anonymity because it distances students from their role in this process, often displaying a lack of commitment or repressive intentions towards the teacher, especially because many have the intention of overseeing the teacher's activity (Hornstein & Edmond, 2017; Spooen et al., 2013; Spooen & Van, 2012), although in recent decades evaluation policies have attempted to focus on this aspect for improvement purposes. Therefore, it does not guarantee high response rates and honest answers from the participants.

Research on questionnaire bias describes multiple reasons why these results are not useful in either teaching development or teacher professional development. This is due to the limited support for following up on observations, the irrelevance of feedback results, a rigid quantitative number which lends itself to associations that fail to reflect the complexity of the evaluated teacher, an insurmountable belief on the part of teachers that it is only applied to comply with the controls required at macro-political level, and finally the impossibility of controlling the differences in evaluations assigned to teachers by one area of knowledge with respect to others, an aspect demonstrated in the study by Uttl and Smibert (2017) in the United States.

In contrast, it seems that lately there are strong convictions that evaluations with students will be more effective when the latter are involved in open, frank dialogues in which they discuss the real aspects that favour or limit the students' professional development in their learning process and those that concern teacher development (Borch et al., 2020). Thus, the most laudable proposal is the triangulated quantitative evaluation with qualitative methods for collecting data from students (Steyn et al., 2019).

In contrast to the arguments against the use of questionnaires with students, this is one of the methods that most influences teachers' own perceptions of their work and that form part of teacher self-evaluations (self-assessment/self-reflection); these methods have a significant impact on the development of teachers' critical thinking skills and the improvement of their ability to innovate (Kalk et al., 2014). The contributions show the methods used in questionnaire validation. However, regardless of the criteria used to determine the validity of the instruments, it is important to consider the factors associated with the level of honesty of students when responding, especially when the teacher is under pressure because these results are used in promotion and continued employment processes in the system.

In some studies the abovementioned questionnaire is complemented by the teaching portfolio. According to Jarauta and Bozu (2013), the teacher portfolio reveals solid evidence on professional pedagogical activity, facilitates the systematization of pedagogical knowledge and in this process of documented reflection on teaching practice itself, genuine learning is created that promotes short-term teacher professional development. The authors insist that when these reflections are shared with peers, their value is taken into account as an effective method of evaluation and professional improvement, especially of new teachers from disciplines other than Pedagogy. This practice of using the teaching portfolio is also defended by Elizalde and Reyes (2008) because it allows teachers to present their reflections and think about the methods, resources and evaluation strategies, and can be technically discussed with peers, hence its usefulness for teacher professional development. These results can be used by teachers for self-diagnosis, for formative purposes by peers and for summative reasons by managers.

In this sense, the value of the examination of the practices and their transformation based on peer evaluation through training programmes was demonstrated. The use of this approach in Australian universities includes observation during a cycle, before, during and after the process. This is important because participants must also observe others (Georgiou et al., 2018; Grainger et al., 2016). Barnard et al. (2014) studied the way in which distributed academic leadership can establish a peer review culture to evaluate teaching, thus facilitating shared and responsible change in people who work together through the combination of initiative and experience. An important element is action research of planning, action, observation, and reflection.

Meanwhile, Yiend et al. (2014) in the United Kingdom studied the individualised peer training process in which the teacher only interacts with an expert and another teacher from the same discipline to obtain better feedback. However, also in the United Kingdom, Thampy et al. (2015) reported on a peer-supported qualitative evaluation with formative objectives. It is worthwhile highlighting the decision to triangulate data to minimize bias produced in students' evaluations and in the negotiation process before starting the process with both the participating teachers and the students. Twelve (12) observations were applied during one year, and then the data were triangulated through semi-structured interviews with students outside the experiment.

Observations can also be performed by recording videos to enable detailed observation of classroom practices, teacher self-analysis and analysis with others to determine improvable aspects. Although the literature does not mention the need for their validation before they can be used, in relation to peer evaluation we agree with Hamel and Viau-Guay (2019) that this approach is sufficiently flexible to adapt to the different teaching styles and levels of development of teachers.

In summary, even though the use of validated and consistent questionnaires and Likert scales presented is widespread, a solution must still be found to overcome the problem of bias that can mask real results and their contribution to teacher professional development. Furthermore, the contributions of the use of observation *in situ* or through videos, as well as teaching portfolios, take for granted their contribution in the self-assessment, reflection and innovation of teachers and peers to strengthen research processes. No references were found in the reviewed literature regarding the use of quasi-experimental or pre- and post-tests, action research as a method to promote the examination of practices and their transformation based on the gathering of data through focal groups and semi-structured interviews. In our opinion, these aspects would complement evaluative research.

Discussion and conclusions

At this point, the following aspects may be highlighted: 1) tensions exist regarding the differences and imbalances in the contributions of each participant to teacher professional development; 2) all the participants evaluated predominantly similar dimensions essentially related to teaching and research, with insufficient attention to a significant number of dimensions, characteristics and skills that contribute to teacher professional development; 3) the methods used were based on segmentation by participant without achieving methodological coherence to ensure a relevant evaluation of aspects of teacher training and professional development. Hence, three points may be highlighted for discussion:

The first point relates to the relevance that direct teacher and peer participation has acquired in the last decade without proposing ways for their individual contributions to be harmoniously integrated in traditional and predominantly student-led evaluation.

The studies mainly highlight individual contributions and the real possibilities of each evaluator to make critical evaluative judgements, principally about teaching and research activity. Student evaluation was analysed in most of the identified studies, with few authors presenting arguments to explain its influence on teacher professional development and weightings that caused imbalances in the contributions of the other participants. In this sense, opinion on this criterion is divided; some authors believe that teacher evaluation does not contribute to professional development due to the bias caused by differences in the disciplines taught by teachers, their gender and the grades they give to students (Aguilar et al., 2014; De Juanas & Beltrán, 2014; Kember & Leung, 2011). This evidence penalises teachers who teach disciplines that are differentiated due to their epistemological nature. Men value work better than women and those who fail students are more likely to obtain much lower ratings.

Other authors consider this to be one of the best sources for obtaining valuable information on teacher performance in class (Márquez & Mother, 2016). However, in our opinion there should be no imbalances in the weightings assigned by other participants in the evaluation. Some authors have also adapted questionnaire models from one country to another, such as the University of Concepción in Chile, which applies the Compound Model adapted from the Spanish context (González et al., 2016). This may be an option for contexts with few studies on the subject, but risky due to cultural differences and the specific level of development of individual teachers.

In the case of peer evaluation of teaching, Iqbal (2014) found that the positive perception of its relevance studied in 30 university teachers in Canada was practically non-existent. The aforementioned results are the consequence of prioritizing the fulfilment of academic productivity and teachers' concerns regarding the inadequate preparation of peers to evaluate others, which fuels distrust in the reliability, results and non-definition of criteria to be taken into account in a genuine evaluation that fosters professional development, especially in the field of teaching. Levander and Riis (2016) were also unable to draw conclusions on the contribution of peers due to their lack of preparation to carry out reviews, thus highlighting the need for peer preparation to ensure the quality of this process.

However, as we advanced with the literature review, we observed a tendency to require teachers to compare their perception of their teaching with what is required in the teaching-learning processes (Akram & Zepeda, 2015; Fernández-Fernández et

al., 2016). Moreover, the opinions of peers/experts are important to foster the development of reflective and critical thinking to reveal strengths and weaknesses in their teaching practice. The use of this space should be established as a priority in formative research groups that examine teaching-learning processes in that area while investigating the area of knowledge. The purpose is not only personal-professional improvement, but also university quality in general (Gómez & Valdés, 2019). Therefore, integrating and balancing the contributions is essential.

The requirement for a participatory approach involving various actors implies the need for integration between the contributions of each actor. For this reason, the manager's role is essential to ensure the continuity and evaluation of teacher development. However, the scarce presence of research highlighting the role of managers in teacher evaluation is significant. There may be two reasons for this: 1) continuous changes at administrative level, requiring teachers to adapt permanently to the new administration, without this representing their professional development in the short and medium term (Rueda, 2008); 2) changes in Higher Education occur too quickly, making managers uncertain of whether they will be able to fulfil more than their established obligations, even if this prevents the strengthening of a real culture of professional growth that transcends the usual desire of managers to protect the image of the institution in question (Aravena & Quiroga, 2018; Rueda, 2008).

This entails the challenge of directing research to chart a theoretical and methodological roadmap in order to ensure coherence in the collaboration of each participant where the manager is the person who evaluates the investigative and professional growth process.

The second point of analysis focuses on achieving consensus in the scientific community with regards to the evaluation of certain dimensions that, to a greater or lesser extent, promote professional development.

In this sense, didactic competence, which is associated with teaching, invariably facilitates consensus on the aspect being evaluated: teaching planning, teaching methodologies, use of ICT equipment and resources, as well as teacher-student communication. This may be associated with the nature of the profession, since the role of teachers has historically been attributed to the establishment of assertive communication mechanisms with students given their traditional role in evaluation.

It is evident from literature that the same dimension is evaluated from different perspectives and to different levels of depth due to the background of each participant. This is natural given the preparation of the different actors participating in the evaluation of a specific dimension such as teaching (evaluated by students and by peers) for example. The subjectivity contributed by each participant in the process and the biases associated with this should not be forgotten.

There is consensus in the literature regarding the evaluation of communication skills, criticism and reflection on the mode of action and educational innovation; these aspects are very often evaluated with greater emphasis by teachers and their peers. The foregoing may stem from the continuous improvement necessary in the teaching process, in which communication with others plays an important role (Pompa & Pérez, 2015).

It is a validated option for the construction of didactic knowledge through shared dialogue between peers, professional discussion, which articulates academic identities and professional knowledge to address the complexity of professional practice that

contributes to growth (Hawes & Troncoso, 2006). The studies document in much more detail the process of collaborative reflection on teaching and feedback on same in the context of academic dialogue involving the entire education community as an essentially valid mechanism not only for the professional development of teachers but also to achieve organisational development characterized by renewal (Garbanzo-Vargas, 2016; Searle et al., 2016; Yan, 2017).

Regardless of the arguments presented on the validity of research linked to teaching practice, research is highly appreciated by teachers themselves. These results may be attributed to the strong pressure they receive to publish impact results in their specialised area of knowledge. In fact, these determine university accreditation indicators and their classification in different rankings. No clear relationship exists between teaching-research and teaching innovation-impact research. They seem to be separate and exclusive aspects with different levels of importance (Jalbert, 2019; Simon et al., 2017).

It is surprising that the frequency with which personal and attitudinal characteristics are evaluated differs from one participant to another. Students evaluate a larger number of characteristics and attitudes in teachers. In our opinion, this may be due to their need to establish closer relationships with teachers, earn their empathy through assertive communication, receive personalised attention in learning and achieve good grades. There is a low presence of these characteristics in the contributions self-evaluated by teachers, perhaps due to the difficulties they encounter in recognising their own characteristics, and they are absent from peer evaluations (Contreras, 2018; Olivera & Costa-Lobo, 2019).

The frequency with which the dimensions of the evaluation are evaluated by the actors involved differs. In some isolated contributions, other dimensions that were present in one and not in others are excluded. In our opinion, their evaluation would also account for the professional development achieved by teachers because they are crucial for the performance of professional functions. Thus, for example, the following aspects were identified in the sphere of teaching: intra- and inter-disciplinary relationships; use of technical language; quality of examples used in learning activities that report linkage to context; feedback and tutoring that make different contributions depending on whether they are given by teachers or students and whether the teacher receives them from a peer. Another interesting finding was that teachers' ability to diagnose learning is not evaluated within the teaching dimension; this is a key aspect from the beginning of the teaching process (Klug et al., 2016).

Few references were found in the reviewed studies to the following aspects: teachers' ability to work with scientific information; the development of projects with students indicating that the teacher had acquired formative research. From the perspective of management competences, only one study described two-way evaluation: one, based on teaching activities, that ensured learning; and the other through activities that improve institutional processes (Ruiz-Corbella & Aguilar-Feijoo, 2017). It is also noteworthy that the key characteristics and attitudes characterising university teaching activities described in the reviewed literature were self-knowledge, perseverance, collaboration, creativity, flexibility and teacher commitment.

This entails a second challenge: focusing research on the in-depth study of the professional functions of teachers, the levels of integration between them, as well as the characteristics and attitudes of teachers that enable an accurate determination of the

dimensional content of the evaluation to foster professional development, as well as the balance weighting of each dimension.

A third point for discussion is the methodological difficulty involved in articulating the dimensions to be evaluated, methods to be used and participants.

The design of instruments that enable more data to be obtained on the different dimensions and participants, as well as the procedures for their triangulation and the preparation of all the participants in the evaluation process, represent challenges in evaluative research. All this together with actions that reduce the factors causing bias in the different instruments and participants. The reviewed literature reveals the criticality of the comprehensive fulfilment of professional functions (teaching, research-engagement and management combined with the personal and attitudinal characteristics of the teacher) and consequently the large number of dimensions to be evaluated, the instrument to be evaluated to be used and the required participants. Understanding this process as one area of evaluative research is helpful for charting a roadmap towards the sustainable improvement of quality in the entire university system (Cancino & Márquez, 2015; Escudero, 2019).

The literature shows that one of the most widely-used methods is the questionnaire in the case of students with great interest in evaluating student satisfaction. Thanks to technological developments, these questionnaires are now completed online, thus casting doubt on their quality given the low response rates, as demonstrated by Spooren and Van (2012). It is also not clear whether students respond honestly for fear of teacher reprisals (Abiodun & Aremu, 2012). And in the case of peer and teacher evaluation, in addition to Likert scales, qualitative methods such as observations, portfolio analysis and interviews are used. In either case, human subjectivity bias may occur and award disproportionate weights and inadequate ratings depending on the educational system in question and the conception of teacher evaluation in each university (Aguilar et al., 2014). As mentioned throughout this study, since several individuals evaluate the teacher and consequently, the use of various methods and instruments is intrinsic to professional development, necessarily making this a complex, mixed investigative process that entails the use of multiple quantitative and qualitative methods (De Diego & Rueda, 2012; Elizalde & Reyes, 2008; Escudero, 2019; Palacios et al., 2019). However, no evidence was found in the reviewed literature to indicate that data triangulation mechanisms make it easier for managers to ensure their effective application.

This entails a third challenge, namely the need for the methodological design to integrate mixed methods and data collection procedures, as well as their triangulated analysis, to identify aspects relevant for teacher training and professional development.

The participatory-collaborative approach can be used as a theoretical concept to transform the professional development practices of teachers and students and foster progress at universities since, by its very essence, it abandons the technocratic and instrumental approach, which is only useful in accountability processes designed to embrace inclusive and contextual approaches that enable participant engagement. It is conceived as an investigative approach based on the social relationships between the evaluators and the interest groups in that context because they experience the same problems and joys. This process can lead to the improvement of the professional quality of teachers through dialogue and professional discussion.

However, in our opinion this research provides an analysis that places great emphasis on the theoretical and practical assumption of collaborative participation in the teacher evaluation process. The researchers' findings undoubtedly contribute to the field of teacher evaluation, but most studies highlight the particularity of each participant while the minority evidence the need to complement the data with other actors and other methods. In view of this situation, literature worldwide has described four cornerstones of the dimensions and methods used in teacher evaluation by each participant, necessary for understanding the contribution of each to teacher and institutional professional development.

Researchers seeking to broaden evaluative research on university teachers may perhaps do so by identifying principles and conditions to create spaces that facilitate reflective debate between the different participants, the evaluated teacher and the manager. They could also search for cross-sectional and longitudinal mixed methodologies: firstly, in periodic evaluations to better integrate the contributions of each participant for the necessary integration between the dimensions of teaching, research-engagement with society and management, with emphasis on the use of technology as an key resource in the international context, and personal and attitudinal characteristics; secondly, to draw attention to the spiral evolutionary process of teacher professional development.

Another stressor worth examining in future research is the limited reference in literature on business management to the procedures for collaborative evaluation practices, possibly due to the absence of an evaluation culture managed through the application of mixed research methods (Suárez et al., 2019) and the inadequate preparation of all the participants and the manager to lead this process. Without mentioning the actual problems of instability of the managers, status and interpersonal relationships that compromise the effectiveness of the contributions.

It is also necessary to take into account the limitations of our study. In our opinion, despite the delimitation by country, no conclusions could be drawn regarding the decision of the researchers from the different regions or countries where the decisions were taken with regard to the dimensions and methods used. Moreover, the scope of the analysis did not include a detailed study of the significance of each dimension for each participant, which may give rise to difficulties in defining contextual indicators and thus undermine the theoretical coherence proposed in the participatory approach, its transparency and credibility. Therefore, future research should focus on conducting a more in-depth study of the perceptions of the dimensions to be evaluated by each participant in order to design participatory and multidimensional evaluation models, and consequently evaluation instruments to gather reliable data about teachers in the context in which they are evaluated with a view to monitoring progress in teacher professional development.

Bibliographic references

- Abiodun, B., & Aremu, O. (2012). A revalidation of students' evaluation of teaching effectiveness rating scale. *The African Symposium: An online journal of the African Educational Research Network*, 12(2), 18-29.
- Aguilar, A. M., Carbonell, A., & Cisneros-Cohernour, E. J. (2014). La evaluación de la docencia en dos universidades públicas latinoamericanas: Cuba y México. *Revista Iberoamericana de Evaluación Educativa*, 7(2e), 65-73.

- Akram, M., & Zepeda, S. J. (2015). Development and Validation of a Teacher Self-assessment Instrument. *Journal of Research and Reflections in Education*, 9(2), 134-148.
- Aravena, F., & Quiroga, M. (2018). Autoetnografía y directivos docentes: una aproximación experiencial a las reformas educativas en Chile. *Revista Electrónica de Investigación Educativa*, 20(2), 113-125. <http://dx.10.24320/redie.2018.20.2.1600>
- Barnard, A., Nash, R., McEvoy, K., Shannon, S., Waters, C., Rochester, S., & Bolt, S. (2014). LeaD-In: LeaD-In: a cultural change model for peer review of teaching in higher education. *Higher Education Research & Development*, 34 (1), 30-44. <http://dx.10.1080 / 07294360.2014.935931>
- Bilbao, A., & Villa, A. (2018). La competencia evaluativa como factor clave en la calidad docente: percepción de las/os maestras/os en la formación inicial. *Profesorado. Revista de Currículum y Formación de Profesorado*, 22(4), 171-195. <http://dx.doi.org/10.30827/profesorado.v22i4.8412>
- Blair, E., & Valdez, K. (2014). Improving higher education practice through student evaluation systems: is the student voice being heard? *Assessment & Evaluation in Higher Education*, 39 (7), 879-894. <http://dx.doi.org/10.1080/02602938.2013.875984>
- Borch, I., Sandvoll, R., & Risor, T. (2020). Discrepancies in purposes of student course evaluations: what does it mean to be "satisfied"? *Educational Assessment, Evaluation and Accountability*, 32, 83-102. <http://dx.10.1007/s11092-020-09315-x>
- Cancino, V. E., & Márquez, T. S. (2015). Evaluación de Desempeño de la Función Académica: Análisis de un Sistema en el Contexto Universitario Chileno. *Formación Universitaria*, 8(3), 35-46. <http://dx.10.4067/S0718-50062015000300005>
- Contreras, G. A. (2018). Retroalimentación por pares en la Docencia Universitaria. Una Alternativa de Evaluación Formativa. *Formación Universitaria*, 11(4), 83-94. <http://dx.10.4067/S0718-50062018000400083>
- Davies, D., Jindal-Snape, D., Digby, R., Howe, A., Collier, C., & Hay, D. (2014). The roles and development needs of teachers to promote creativity: A systematic review of literatura. *Teaching and Teacher Education*, 41, 34-41. <http://dx.10.1016/j.tate.2014.03.003>
- 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*, 7(2), 479-515. <http://dx.0.20511/pyr2019.v7n2.255>
- De Juanas, A., & Beltrán, J. A. (2014). Valoraciones de los estudiantes de ciencias de la educación sobre la calidad de la docencia universitaria. *Educación XX1*, 17(1), 59-82. <http://dx.10.5944/educxx1.17.1.10705>
- Dios, I., Calmaestra, J., & Rodríguez-Hidalgo, A. J. (2018). Validación de la escala de competencias docentes organizacionales y didácticas para educadores. *Revista Mexicana de Investigación Educativa*, 23(76), 281-302.
- Donado, A. C., Zepa, C. E., & Ruiz, B. L. (2018). Academic Engagement, Academic Achievement, and Teacher Quality According to Gender: A Study with University Students from the Colombian Caribbean. *New directions for teaching and learning*, 156, 49-56. <http://dx.doi.org/10.1002/tl.20316>
- Eizagirre, A., Altuna, J., & Fernández, I. (2017). Prácticas de éxito en el desarrollo de competencias transversales en centros de Formación Profesional del País Vasco. *Revista Española de Pedagogía*, 75(267), 293-308. <http://dx.doi.org/10.22550/REP75-2-2017-7>

- Elizalde, L., & Reyes, R. (2008). Elementos clave para la evaluación del desempeño de los docentes. *Revista Electrónica de Investigación Educativa*, 10(Esp), 1-13.
- Escobar, M. B. (2015). Influencia de la interacción alumno-docente en el proceso enseñanza-aprendizaje. *AAKAT: Revista de Tecnología y Sociedad*, 0(8).
- Escudero, T. (2019). Evaluación del profesorado como camino directo hacia la mejora de la calidad educativa. *Revista de Investigación Educativa*, 37(1), 15-37. <http://dx.10.6018/rie.37.1.342521>
- Fernández, F. (2002). El análisis de contenido como ayuda metodológica para la investigación. *Revista de Ciencias Sociales, II* (96), 35-53.
- Fernández-Fernández, S., Arias-Blanco, J. M., Fernández-Alonso, R., Burguera-Condon, J., & Fernández-Raigoso, M. (2016). Pensamiento reflexivo e investigador en educación. Aspectos a tener en cuenta en la formación del profesorado. *RELIEVE*, 22(2), 1-16. <http://dx.doi.org/10.7203/relieve.22.2.8425>
- Ferreira, I., Urrútia, G., & Alonso-Coello, P. (2011). Systematic Reviews and Meta-Analysis: Scientific Rationale and Interpretation. *Revista Española de Cardiología* 64(8), 688-696. <http://dx.doi:1016/j.recesp.2011.03.029>
- Garbanzo-Vargas, G. M. (2016). Desarrollo organizacional y los procesos de cambio en las instituciones educativas, un reto de la gestión de la educación. *Educación*, 40(1), 67-87. <http://dx.10.15517/revedu.v40i1.22534>
- Georgiou, H., Sharma, M., & Ling, A. (2018). Peer review of teaching: What features matter? A case study within STEM faculties. *Innovation in Education and Teaching International*, 55(2), 190-200. <http://dx.10.1080/14703297.2017.1342557>
- Gómez, L. F., & Valdés, M. G. (2019). La evaluación del desempeño docente en la educación superior. *Propósitos y Representaciones*, 7(2), 479-515. <http://dx.10.20511/pyr2019.v7n2.255>
- González, I., López, A., & Kroyer, N. (2016). Claves de Compround para la redefinición del modelo de evaluación de la calidad docente en la Universidad de Concepción. *Estudios Pedagógicos XLII*, (4), 69-85. <http://dx.10.4067/S0718-07052016000500005>
- Grainger, P., Crimmins, G., Burton, K., & Oprescu, F. (2016). Peer review of teaching (PRoT) in higher education – a practitioner's reflection. *Reflective Practice*, 17(5), 523-534. <http://dx.10.1080/14623943.2016.1146581>
- Hamel, C., & Viau-Guay, A. (2019). Using video to support teachers' reflective practice: A literatura review. *Cogent Education*, 6(1), 1-14. <http://dx.10.1080/2331186X.2019.1673689>.
- Hawes, G., & Troncoso, K. (2006). A propósito de evaluación por pares: la necesidad de sistematizar la evaluación y las prácticas docentes. *Perspectiva Educacional*, (48), 59-72.
- Hornstein, H., & Edmond, H. F. (2017). Student evaluations of teaching are an inadequate assessment tool for evaluating faculty performance. *Cogent Education*, 4, 1-8. <http://dx.10.1080/2331186X.2017.1304016>
- Hortigüela, D., Ausín, V., Delgado, V., & Abella, V. (2017). Análisis de la importancia de los criterios de evaluación y el reconocimiento académico docente universitario como indicadores de la calidad educativa en España. *Revista de la Educación Superior*, 46(181), 75-87. <http://dx.10.1016/j.resu.2016.10.002>
- Iqbal, I. (2014). Don't tell it like it is: Preserving Collegiality in the Summative Peer Review of Teaching. *Canadian Journal of Higher Education*, 44(1), 108-124.

- Isla-Díaz, R., Marrero-Hernández, H., Hess-Medler, S., Soriano, M., Acosta-Rodríguez, S., Pérez-Montevedre, M. V., & Blanco-Freijo, M. (2018). Una mirada longitudinal: ¿Es el "Docentia" útil para la evaluación del profesorado universitario? *RELIEVE*, 24(2), 1-21. <http://dx.10.7203/relieve.24.2.12142>.
- Jalbert, T. (2019). Relationships between business faculty teaching and research ratings. *Research in Higher Education Journal*, 36, 1-20.
- Jara, N. P., & Díaz-López, M. M. (2017). Políticas de evaluación del desempeño del docente universitario, mito o realidad. *Educación Médica Superior*, 31(2).
- Jarauta, B., & Bozu, Z. (2013). Portafolio docente y formación pedagógica inicial del profesorado universitario. Un estudio cualitativo en la Universidad de Barcelona. *Educación XX1*, 16 (2), 343-362. <http://dx.doi.org/10.5944/educxx1.2.16.10345>
- Jornet, J. M., González-Such, J., Suárez, J. M., & Perales, M. J. (2011). Diseños de evaluación de competencias: consideraciones acerca de los estándares en el dominio de las competencias. *Bordón*, 63 (1), 125-145.
- Kalk, K., Luik, P., Taimalu, M., & Täht, K. (2014). Validity and reliability of two instruments to measure reflection: a confirmatory study. *Trames*, 18(2), 121-134. <http://dx.10.3176/tr.2014.2.02>
- Kember, D., & Leung, Y. P. (2011). Disciplinary differences in student ratings of teaching quality. *Research in Higher Education*, 52(3), 278-99. <http://dx.10.1007/s11162-0109194-z>
- Klug, J., Bruder, S., & Schmitz, B. (2016). Which variables predict teachers' diagnostic competence when diagnosing students' learning behavior at different stages of a teacher's career? *Teachers and Teaching: Theory and Practice*, 22(4), 461-484. <http://dx.10.1080/13540602.2015.1082729>
- Levander, S., & Riis, U. (2016). Assessing educational expertise in academic faculty promotion. *Nordic Journal of Studies in Educational Policy*, 2(3), 1-13. <http://dx.10.3402/nstep.v2.33759>
- Lizasoain-Hernández, L., Etxebarria-Murgiondo, J., & Lukas-Mujjika, J. F. (2017). Propuesta de un nuevo cuestionario de evaluación de los profesores de la Universidad del País Vasco. Estudio psicométrico, dimensional y diferencial. *RELIEVE*, 23(2). <http://dx.10.7203/relieve.23.2.10436>
- Loredo, J., Romero, R., & Inda, P. (2008). Comprensión de la práctica y la evaluación docente en el posgrado a partir de la percepción de los profesores. *Revista Electrónica de Investigación Educativa*, 10 (Especial), 1-16.
- Luna, E., & Reyes, E. (2015). Validación de constructo de un cuestionario de evaluación de la competencia docente. *Revista Electrónica de Investigación Educativa*, 17(3), 13-27.
- Márquez, L., & Madueño, M. L. (2016). Propiedades psicométricas de un instrumento para apoyar el proceso de evaluación del docente universitario. *Revista Electrónica de Investigación Educativa*, 18(2), 53-61.
- Mas, O. (2014). Las competencias investigadoras del profesor universitario: la percepción del propio protagonista, de los alumnos y de los expertos. *Profesorado. Revista de currículo y formación del profesorado*, 18(3), 255-273.
- Moreira, L. M., & Santos, M. Á. (2016). Evaluando la enseñanza en la Educación Superior: percepciones de docentes y discentes. *Revista Electrónica de Investigación educativa*, 18(3), 19-36.

- Moreno-Murcia, J. A., Silveira, Y., & Belando, N. (2015). Questionnaire evaluating teaching competencies in the university environment. Evaluation of teaching competencies in the university. *New approaches in educational research*, 4(1), 54-61. <http://dx.10.7821/naer.2015.1.106>
- Navarro, C., & Ramírez, M. S. (2018). Mapeo sistemático de la literatura sobre evaluación docente (2013-2017). *Educación e Investigación*, 44.(e185677), 1-23. <http://dx.doi.org/10.1590/S1678>
- Nyame, F., Alhassan, I., & Alhassan, A. (2019). Multivariate Analysis of Students "Perception of the Impact of Lecturers" Ranks on their Performance at the Faculty of Mathematical Sciences. *Higher Education Studies*, 9(1), 53-62. <http://dx.doi.org/10.5539/hes.v9n1p53>
- Olivera, S., & Costa-Lobo, C. (2019). Evaluate the pedagogical practice of the teachers of Higher Education: A Proposal. *REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 17(1), 39-54. <http://dx.doi.org/10.15366/reice2019.17.1.003>
- Pompa, Y. C., & Pérez, I. A. (2015). La competencia comunicativa en la labor pedagógica. *Universidad y Sociedad*, 7(2), 160-167.
- Rueda, M. (2008). La evaluación del desempeño docente en la universidad. *Revista Electrónica de Investigación Educativa*, (Especial), 1-15.
- Ruiz, C., Mas, O., Tejada, J., & Navío, A. (2008). Funciones y escenarios de actuación del profesor universitario. Apuntes para la definición del perfil basado en competencia. *Revista de la Educación Superior*, 2 (146), 115-132.
- Ruiz-Corbella, M., & Aguilar-Feijoo, R. (2017). Competencias del profesor universitario: elaboración y validación de un cuestionario de autoevaluación. *Revista Iberoamericana de Educación Superior*, VIII (21), 37-65.
- Sammaknejad, A., & Marzban, A. (2016). An Analysis of Teachers' Self-reflection on-Classroom Management. *Theory and Practice in Language Studies*, 6(1), 84-89. <http://dx.10.17507/tpis.0601.11>
- Searle, M. J., Merchant, S., Chalas, A., & Lam, Y. L. (2016). A Case Study of the Guiding Principles for Collaborative Approaches to Evaluation in a Developmental Evaluation Context. *Canadian Journal of Program Evaluation*, 31(Special Issue), 350-373. <http://dx.10.3138/cjpe.328>
- Serrano, R., Macías, W., Rodríguez, K., & Amor, M. (2019). Validating a scale for measuring teachers' expectations about generic competences in higher education: The Ecuadorian case". *Journal of Applied Research in Higher Education*, 11(3), 439-451. <http://dx.10.1108/JARHE-09-2018-0192>
- Simon, C., Dimovski, V., & Zarman, M. G. (2017). Research, teaching and performance evaluation in academia: the salience of quality. *Studies in Higher Education*, 42 (8), 1455-1473. <http://dx.10.1080/03075079.2015.1104659>
- Spooren, P., & Van, F. (2012). Who participates (not)? A non-response analysis on students' evaluations of teaching. *Procedia - Social and Behavioral Sciences*, 69, 990-996. <http://dx.10.1016/j.sbspro.2012.12.025>
- Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the Validity of Student Education of Teaching: The State of the Art. *Review of Educational Research*, 83 (4), 598-642. <http://dx.10.3102/0034654313496870>
- Steyn, C., Davies, C., & Sambo, A. (2019). Eliciting student feedback for course development: the application of a qualitative course evaluation tool among business re-

- search students. *Assessment & Evaluation in Higher Education*, 44(1), 11–24. <http://dx.10.1080/02602938.2018.1466266>.
- Suárez, N., Mena, D., Gómez, V., & Fernández, A. I. (2019). *La formación del profesorado en Iberoamérica. Tendencias, reflexiones y experiencias*. Quito, Ecuador: Universidad Tecnológica Indoamérica.
- Suárez, N., Palacios, D. A., Delgado, K. E., & Pérez, I. C. (2019). Complejidades del desarrollo profesional universitario y claves metodológicas mixtas para su análisis. *Revista Cubana de Medicina Militar*, 48(2 sup), 389-409.
- Suárez, N., Yaguana P. C., & Gómez, V. (2018). Cualidades, habilidades y competencias docentes. Claves para innovar y liderar procesos en educación. In N. Suárez (Ed.), *Investigación e Innovación en Educación* (pp. 159-184). Quito, Ecuador: Universidad Indoamérica.
- Thampy, H., Bourke, M., & Naran, P. (2015). Peer-supported review of teaching: an evaluation. *Education for primary care*, 26(5), 306–310. <http://dx.10.1080/14739879.2015.1079020>
- Thomas, S., Chie, Q. T., Abraham, M., Jalarajan Raj, S., & Beh, L. S. (2014). A Qualitative Review of Literature on Peer Review of Teaching in Higher Education. *Review of Educational Research*, 84(1), 112-159. <http://dx.10.3102/0034654313499617>
- Tyunnikov, Y. S. (2016). Interrelation of Evaluation and Self-Evaluation in the Diagnostic Procedures to Assess Teachers' Readiness for Innovation. *European Journal of Contemporary Education*, 16(2), 248-256. <http://dx.10.13187/ejced.2016.16.248>
- Uttl, B., & Smibert, D. (2017). Student evaluations of teaching: teaching quantitative courses can be hazardous to one's career. *PeerJ*, 5(5), 1-13. <http://dx.10.7717/peerj.3299>
- Wellein, M. G., Ragucci, K. R., & Lapointe, M. (2009). A peer review process for classroom teaching. *American Journal of Pharmaceutical Education*, 73(5). <http://dx.10.5688 / aj730579>
- Xu, Y. (2012). Developing a comprehensive teaching evaluation system for foundation courses with enhanced validity and reliability. *Educational Technology Research and Development*, 60(5), 821-837. <http://dx.10.1007/s11423-012-9240-y>
- Yan, Ch. (2017). The Oral History of Evaluation: An Interview with Lyn Shulha. *Canadian Journal of Program Evaluation*, (Número Especial), 397–408. <http://dx.10.3138/cjpe.32>
- Yiend, Y., Weller, S., & Kinchin, L. (2014). Peer observation of teaching: The interaction between peer review and developmental models of practice. *Journal of Further and Higher Education*, 38(4), 465-484. <http://dx.10.1080/0309877X.2012.726967>.