Regular Article

Educational Policy
1–27

DOI: 10.1177/0895904820951123
journals.sagepub.com/home/epx
SAGE

A Model for Developing an Academic Activity Index for Higher Education Instructors Based on Composite Indicators

María del Carmen Bas[1] and José Miguel Carot[2]

Abstract

The assessment of the teacher performance is a subject of criticism due to the lack of a well-establish methodology. This study develops an overall score to measure the dimensions that encompass the academic activities. To that end, a Benefit-of-the-doubt model is proposed. The advantage of this technique is the flexibility in the weights, so that the model selects for each teacher the most favourable set of weights. Furthermore, the paper proposes the barycentric coordinate system as a method to classify the teachers in clusters depending on their contribution to the dimensions. A specific pie chart has been proposed as an efficient way to report the contribution of the teachers to the dimensions and the overall teacher performance.

Keywords

benefit-of-the doubt, academic activity index, teacher profiles, barycentric coordinate system, composite indicators

[1]Departamento de Matemáticas para la Economía y la Empresa, Universitat de València, Valencia, Spain
[2]Centre for Quality and Change Management, Universidad Politécnica de Valencia, Valencia, Spain

Corresponding Author:
María del Carmen Bas, Departamento de Matemáticas para la Economía y la Empresa, Universitat de València, Av/ Tarongers s/n, València, 46022, Spain.
Email: maria.c.bas@uv.es

Introduction

In the past years, universities have undertaken various changes in the academic activities of teaching, research and university management, which have been conditioned by the demands of society. The assessment of the teachers' activity is a subject of study by many higher education institutions in different countries. However, despite the growing interest of higher education systems for having measurement tools to evaluate the activity of teachers, the issue is still under constant discrepancies. The main reasons that promote these discussions are the lack of a common framework to evaluate the teachers' activities, and the lack of a well-established methodology to justify the assessment process.

Which activities should be included in the higher education teachers' assessment? What is the definition of an "effective" teacher? How can weight and aggregate be applied to the different activities of higher education teachers? What is the potential tool to evaluate the teachers' activities? These are some of the questions that are constantly raised in the assessment process of teachers' activities.

Shifting missions and activities of institutions striving to emulate those that are highly ranked have affected the ways in which teachers engage in academic work. The expansion of teacher roles as one of the most important issues facing teachers on their campuses resulted in an intensification of tasks expected of teachers and pressure to keep up with changes in teaching and research directions, to find ways to support a more diversified student body, and uphold increased research productivity (Baker, 2017).

Research on professional growth in academic environments suggests that opportunities to learn and grow in teaching will likely influence teacher productivity (Campbell & O'Meara 2013). In research led by O'Meara et al. (2017), teacher's professional growth is evaluated through the faculty scholarly learning, defined as the "process in which teachers increase subject matter knowledge and skills" to be a good teacher. What teachers learn as scholars varies based on their interests; work roles (teaching, research, internal service and outreach); the groups with which they interact and the organizational contexts. The teacher's professional growth has significant implications for the entire system of faculty recruitment, assessment, support, and reward with regard to engagement.

Schuster and Finkelstein (2006) affirm that beyond the total work effort that teachers devote to teaching, research, and other institutional responsibilities, is the matter of how they distribute their time among these competing responsibilities. The teacher surveys obtained in their research offer two lenses through which to view faculty division of labour. Although they

provide self-reports of actual division of labour, they also give insights into the preferred division of labour, that is, the value that teachers place on these activities both personally and professionally.

In most of the previous studies, the authors applied descriptive statistics to survey methods to explore all the concepts that could influence teacher productivity. Alternatively, this study is based on the construction of composite indicators, which are tools that allow us to evaluate complex multidimensional concepts such as the assessment of the teachers' activity. A composite indicator is defined as a mathematical aggregation of a set of simple indicators. However, the construction of a composite indicator involves subjective steps that should be assessed carefully (Saltelli, 2007). For that reason, there exists different opinion about the use of composite indicators. On the one hand "the aggregators think that a composite indicator is a summary statistic that can indeed capture reality and is meaningful, and that stressing the bottom line is extremely useful in garnering media interest and hence the attention of policy makers. On the other hand, the non-aggregators think one should stop once an appropriate set of indicators has been created and not take the further step of producing a composite index" (Sharpe, 2004).

Therefore, the aggregation step is a controversial topic in the development of a composite indicator where weights may represent the relative importance of each simple indicator. However, how important are the teaching, research and university management activities in the assessment of the overall level of effectiveness of the higher education teachers?

Arneson (1990), Fleurbaey (1992) and Rawls (1999), among others, presented what is known by "indexing dilemma" in the social choice literature. The indexing dilemma discusses the difference between imposing a fixed set of weights to all observations, independent of the diverse personal preferences or accounting for each observation's individual preferences by using individual weights.

There are many studies that critique the idea of assigning a fixed and equal set of weights for all teachers (Avital and Collopy, 2001; Emery et al., 2003; Marsh and Roche, 1994; Pritchard et al., 1998;). The authors pointed that the activities of the teachers are too complex to summarise them with an equal/fixed weighted average. Furthermore, even though the teaching, research and university management activities are very important to measure the effectiveness of the teachers there exists diverging beliefs about what is meant by "effective teacher."

The theoretical framework defined by Blackburn and Lawrence (1995) explains teachers' behaviour in three key areas, research, teaching and service. To do that, the model includes socio-demographic characteristics, career variables describing the career path, self-knowledge variables (contain

self-perceived beliefs, attitudes, and values, such as one's efficacy as a researcher and one's level of ambition and persistence) and social knowledge variables (indicate how the teachers perceive the environment). After applying the model to a representative sample, the results show that the self-knowledge and social knowledge variables were found to be more important than the usual socio-demographic and career variables. Therefore, teachers are motivated by both self-knowledge (their intrinsic commitment and sense of personal agency) and social knowledge (their sense of the kinds of learning and contributions the institutions and their colleagues most value).

In the same line, following the conceptual framework for studying professional growth defined by O'Meara et al. (2008, 2017), the authors describe faculty growth as (i) on-going and in a constant state of becoming as opposed to being fixed, (ii) a process that is facilitated by external environments but that also must be viewed in terms of what individuals themselves want and need as developing persons, and (iii) set in a specific sociocultural and personal context relative to teachers' identities, roles, and work. The conceptual framework explores teachers' professional development through four aspects: (i) learning (given the different nature of their disciplines, teachers carry out their learning in their many different work roles -teaching, research, internal service and outreach- in different ways), (ii) agency (agency speaks to a feature of the faculty person in the faculty role as she or he strives to construct the contexts of her or his own learning and development in professional and intellectual ways), (iii) professional relationships (considering different work roles, teachers have different interactions with students, colleagues, and the public) and, (iv) commitments (ability to act on and form). In conclusion, opportunities for learning, a sense of agency to plan for and seize such opportunities, professional relationships, and a sense of community or connection, anchored in commitment, all shape the different degree to which teachers make contributions to the different work roles and to students, colleagues, institutions, and society. This framework for faculty growth has significant implications for the entire system of faculty recruitment, support, and reward with regard to engagement.

In both frameworks, faculty growth is a continuous process that is driven both by what individuals themselves want and need and by the specific sociocultural, institutional, and personal context in which their identities, roles, and work are defined. These theories suggest the idea of using individual specific weights according to the teacher profile, instead of using fixed weights. To that end, a Benefit-of-the-Doubt methodology (BoD) (Cherchye et al., 2007; Melyn and Moesen, 1991; Rogge, 2011b), which is a version of the non-parametric technique Data Envelopment Analysis, is needed to evaluate the productivity of teachers.

Several studies applied the non-parametric Data Envelopment Analysis model to evaluate the higher education teaching and research activities (Bana e Costa & Oliveira, 2011; Beasley, 1995; De Witte & Rogge, 2011; De Witte et al., 2013; Giménez & Martínez, 2006; Pina & Torres, 1995; Rogge, 2011a; Tomkins & Green, 1988). There exist other studies that have proposed the development of a model focusing only on research activities of the university teachers (Johnes & Johnes, 2009) or a model only based on the teaching activities (Ramón et al., 2010).

The aim of this paper is to propose a model to measure the teachers' activities according to the teacher profile. The teachers' activities proposed in the model include those of the teaching mission, activities and the research mission, and the management activities that support these two main missions of the university teachers. To that end, the Benefit-of-the-Doubt methodology (BoD) was applied.

Benefit-of-the-Doubt model for Teacher Evaluation

Theoretical Conceptual Framework for the Academic Activity Index

The aim of the Academic Activity Index (AAI) is to measure in a single value the different activities of teaching, research and management of teachers in a Spanish university, the Universitat Politècnica de València (UPV). These activities are in turn multidimensional concepts. Therefore, the AAI is based on a set of composite indicators previously defined and accepted by the university case study.

The Teaching Activity Index (TAI) aims to measure the teachers' activities considering their professional categories and their contribution to achieve the objectives of the degrees. Appendix A presents the indicators evaluated in the TAI index. For comprehensive information about the weights assigned to each indicator refer to report developed by the university (Bas, 2014; Bas et al., 2017; Cano et al., 2009; Vicerrectorado de Calidad y Evaluación de la Actividad Académica, 2014).

The intention of the Research Activity Index (RAI) is to measure the work and results in research, technological or artistic development, and knowledge transfer and innovation of the teachers. Appendix A presents the indicators evaluated in RAI index. For detailed information about the indicators' weights measured in the RAI see the report developed by the university (Vicerrectorado de Investigación, Innovación y Transferencia, 2008).

Finally, the Management Activity Index (MAI) attempts to measure the management activities of the teachers, which supports the Teaching and Research missions of the university teachers. The main activities that constitute this index are the participation in committee meetings and in positions that manage the university (head of department, dean or vice-dean of faculty, vice-chancellor, chancellor, etc.). In general, MAI measure the quantity of time dedicated to participating in the teachers' activities that support the teaching, research and management missions of the institution. Appendix A presents the indicators evaluated in the MAI index.

The AAI proposed encompasses the three dimensions mentioned. The CI defined for each dimension was accepted by the university case study by fixing a weight for each activity. In that case, it was easy to reach a consensus about which activity should weight more than others in the final CI. However, it is not easy to find a fixed weight for the TAI, RAI and MAI index. The disagreement among teachers and experts on the specific weighting scheme used to aggregate the dimensions and, on the difficulty, to define what "effective teacher" means, is an indicative of using individual specific weights according to the teacher profile.

Due to the nature of these activities of the three dimensions and the way in which the teachers carry them out, a linear compensatory aggregation procedure is proposed. This is a very simple method that allows the low values of some indicators to compensate for higher values in other indicators.

The AAI is defined as follows:

$$AAI_c = w_1 TAI_c + w_2 RAI_c + w_3 MAI_c$$

where w_i are the weight of each dimension and c is the evaluated teacher.

According to the different types of contracts with the university case study, there are diverse university positions or categories, which should be considering when analysing whether or not the AAI is able to gather the peculiarities of each case. Appendix B shows a description of the responsibilities of each category.

- Non-PhD assistant lecturers (A)
- Non-PhD lecturers (B)
- PhD assistant lecturers (C)
- PhD lecturers (D)
- Non-PhD Senior Lecturers (E)
- Senior Lecturers (F)
- Full Professors (G)

The "Benefit-of-the-Doubt" Methodology

The Benefit-of-the-Doubt method is a version of the Data Envelopment Analysis model (DEA) developed by Charnes et al. (1978), which is a widely used technique for measuring the efficiency performance of organisations, countries, units, etc. The DEA is a linear model technique for measuring the efficiency performance of a set of comparable observations that use inputs to produce outputs with no information on the weights of inputs and outputs and/or no knowledge about the functional form of the production function.

The DEA-method is a widely used tool for constructing "Benefit-of-the-Doubt" composite indicators. Melyn and Moesen (1991) were the pioneers who proposed the use of DEA-model in the context of CIs (known as BoD-model) to evaluate macroeconomic performance. It is evident the similarity between the DEA approach and the construction of composite indicators, in which simple indicators are known and their weights are unknown. The BoD-model only differs from the DEA-model as it exclusively focuses on aggregating outputs considering a "dummy input" vector that is unity.

The idea of the BoD-model is to select endogenously the weights for the indicators using the data set available, unknowing, which are the correct weights for each simple indicator. In the context of the study, the BoD-model assigns teachers' performance in a relative perspective by comparing them to each other and looking for the set of weights that maximise the impact of academic criteria of relative strength and minimise the influence of items of relative weaken. The BoD-model selects for each teacher the most favourable weights. Therefore, BoD-model grants each teacher the benefit-of-the-doubt when it assigns the set of weights in the evaluation of his/her performance considering his/her teacher profile.

In formal terms, the BoD-model could be formulated as follows:

$$CI_c^* = \max_{(w_{c,1},\ldots,w_{c,q})} \frac{\sum_{i=1}^{q} w_{c,i} I_{c,i}}{\max_{(I_{1,j},\ldots,I_{n,j})} \sum_{i=1}^{q} w_{c,i} I_{j,i}}$$

$$s.t.$$

$$\sum_{i=1}^{q} w_{c,i} I_{j,i} \leq 1 \forall j = 1,\ldots,c,\ldots,n$$

$$w_{c,i} \geq 0 \quad i = 1,\ldots,q$$

In this notation, n represents the number of teachers, CI_c^* is the BoD-optimal score for each teacher c, $I_{c,i}$ is the score of the indicator i for the teacher c, $w_{c,i}$ is the optimal BoD-weight assigned to the indicator i for the c, and q is the number of indicators evaluated.

Note that the ratio format of the objective function reveals the benchmarking objective of the BoD-model. In this way, a $CI_c^* = 1$ implies a global level of teacher's performance, which means that no other receives a higher CI_c^* using the BoD-optimal weights of the evaluated performance c. Otherwise, a $CI_c^* < 1$ refers to a lower global level of teacher's performance, which means that exists at least one who is evaluated with a higher CI_c^* when applying the BoD-weights, $w_{c,i}$, for the evaluated performance c.

Since, by constructing the benchmark performance of teachers obtains a BoD-optimal score of 1, the maximisation problem can be written in a linear form as follows:

$$CI_c^* = \max_{(w_{c,1},\ldots,w_{c,q})} \sum_{i=1}^{q} w_{c,i} I_{c,i}$$

The BoD-model has many conceptual and methodological advantages (Cherchye et al., 2007; Rogge, 2011b): (i) the model is unit invariance, which means that the value of the CI is independent of the units of measurement of the indicators; (ii) the model allows flexibility in the weighting; (iii) the weights are determined to ensure that an optimal score is realised for each unit; (iv) it is easy to incorporate information on appropriate weights in case it is available; and, finally, (v) the BoD-weights reflect the indicators of relative strength and weakness in the evaluated unit, and it is easy to compute the contribution (known as pie-shares, $P_{c,i}^* = w_{c,i} I_{c,i}$) of each indicator to the BoD-optimal score.

The BoD-model clearly marks a deviation from the equal/fixed weighting scheme, widely criticised by many authors. As pointed out by Rogge (2011a), the large flexibility in weight choice could be an advantage because it enables teachers to put themselves in the best possible light relative to their colleagues. Therefore, the model assigns to each teacher an optimal set of weights to obtain the maximum value of the CI. Any other weighting scheme than the one specified by BoD-model would worsen the CI value. However, this flexibility also carries some disadvantages as it may allow a teacher to appear as a brilliant performer in a manner that is hard to justify by the experts. Furthermore, it could happen that the use of a full-flexibility BoD-model ignores some indicators in the assessment of a performance while overemphasising others. Roll, Cook and Golany (1991) stated that it is very difficult to accept the ignorance of indicators in the CI score after having a consensus on the set of relevant indicators to evaluate the performance of the university teachers. Consequently, the is a risk of obtaining CIs scores based only on a few of the indicators evaluated. To avoid unrealistic additional weight, restrictions based on the experts' opinion are recommended in the

BoD and DEA literature (Allen et al., 1997; Cherchye et al., 2007; Cherchye et al., 2008; Thanassoulis et al., 2004).

In the BoD and DEA literature, different weight restrictions have been defined (e.g., (Roll and Golany, 1993; Thanassoulis and Allen, 1998; Thompson et al., 1990; Wong and Beasley, 1990)). In this paper we have used the ones proposed by Wong and Beasley (1990), because they do not apply to the pure BoD-weights, $w_{c,i}$, but to the product of the pure BoD weight and the score of the indicator i for the c, $I_{c,i}$, as follows:

$$\alpha_i \leq \frac{w_{c,i} I_{c,i}}{\sum_{i=1}^{q} w_{c,i} I_{c,i}} \leq \beta_i$$

The restrictions on the product $w_{c,i} I_{c,i}$ called "pie-share" have the advantages of interpretation and unit invariance (Cherchye et al., 2007). The pie-shares indicates how each indicator contributes to the overall value of the CI and this may be attractive to detect indicators that need to be improved. This interpretation is easier for the experts and stakeholders to quantify their opinion in terms of percentage value. Moreover, the resulting CI remains invariant to the units of measurement.

Results

Data Collection Process

Most of the data for developing composite indicators TAI, RAI and MAI were collected from institutional databases of the Universitat Politècnica de València. Teachers in a personalized platform should implement the indicators, which are not in institutional bases. The evaluation procedure of the teachers' activity is carried out every academic year in which teachers have personalized access to the platform. Teachers add information in the platform that is not included in the institutional databases in order to complete the information required in the development of the AAI. Teachers are required to validate the information added in the platform through appropriate certificates. All the indicators defined in TAI, RAI and MAI (see Appendix A) are observed indicators except "Teaching materials" and "Student opinion survey" indicators, which are based on survey methods.

Weighting and Aggregating the TAI, RAI and MAI Index

The total population of teachers of the Universitat Politècnica de València (2,020 teachers) form the database of study containing the three composite indicators (TAI, RAI, MAI), which are the base for the development of the

Table 1. Descriptive Statistics for the Composite Indicators.

	MEAN	ST. DEV.	MIN	MAX
TAI	47.11	10.64	9.41	68.50
RAI	49.69	53.63	0.00	208.00
MAI	2.00	3.55	0.00	20.00

AAI. The descriptive statistics (Table 1) show the heterogeneous range of variation among indicators, which are measured, in different units.

Subsequently, we had used the weight restrictions introduced by Wong and Beasley (1990). The opinion of thirteen expert teachers of the Teaching Evaluation Committee (TEC) was required to quantify the importance of the teaching, research and management dimensions using the "budget allocation" technique which is a participatory technique that require each expert to allocate 100 points among the dimensions; based on the desired importance they gave them. This process provided bounds for each dimension. Then, a final consensus is reached (among expert teachers) on the desired bounds to be allocated to each dimension (Nardo et al., 2005). The restrictions allow a maximum contribution in the teaching and research dimensions of 95% and a minimum one of 1%, considering the different profiles of contract. Furthermore, the maximum contribution for the management activities can be 80% and there is not a minimum one due to that the junior assistant teachers do not enough experience to do management activities and it is not an obligation.

The BoD model proposed to obtain an AAI for the teachers is as follows:

$$AAI_c^* = \max_{(w_{c,1},\ldots,w_{c,q})} \sum_{i=1}^{3} w_{c,i} I_{c,i}$$

$$s.t.$$

$$\sum_{i=1}^{3} w_{c,i} I_{j,i} \leq 1 \forall j = 1,\ldots,c,\ldots,2020$$

$$0.01 \leq \frac{w_{c,1} I_{c,1}}{w_{c,1} I_{c,1} + w_{c,2} I_{c,2} + w_{c,3} y_{c,3}} \leq 0.95$$

$$0.01 \leq \frac{w_{c,2} I_{c,2}}{w_{c,1} I_{c,1} + w_{c,2} I_{c,2} + w_{c,3} y_{c,3}} \leq 0.95$$

$$0 \leq \frac{w_{c,3} I_{c,3}}{w_{c,1} I_{c,1} + w_{c,2} I_{c,2} + w_{c,3} y_{c,3}} \leq 0.8$$

$$w_{c,i} \geq 0 \qquad i = 1,\ldots,3$$

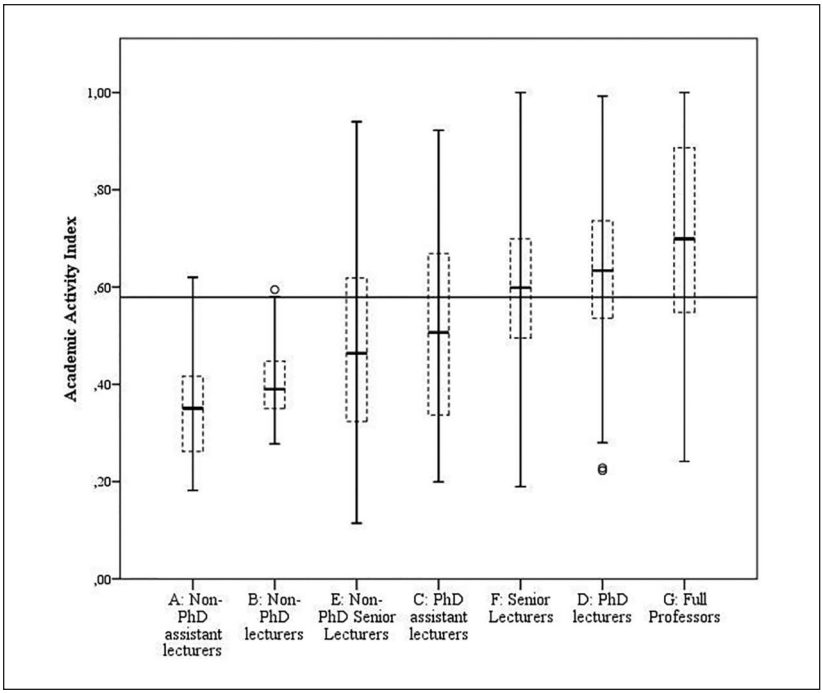

Figure 1. AAI by teachers' categories.

The application of this model allows for a set of individualised weights for each teacher and, therefore, a value in their AAI.

Figure 1 shows the Box and Whisker representation for the different teachers' categories. Note that the A, B, E and C categories have a mean AAI lower than the global mean AAI obtained including all the categories. By contrast, F, D and G categories, which correspond to teachers with a much more consolidated academic career, have higher mean value in AAI than the global mean. This result is consistent with the research of Winslow (2010) which observed that time at university is the most important resource teachers need to accomplish career goals.

However, there is a high variability in the AAI of each category, and this result suggests that the composite indicator penalize or awards teachers depending on the academic, research and management activities without considering the teachers' category.

Visualising Teacher Profiles in Barycentric Coordinate System

To have a clear vision of how AAI represents the activity of teachers in each one of the three dimensions considered, a barycentric coordinate system has been used.

In this study, the cartesian coordinates of the dimension's contribution (α,β,γ), which correspond to the pie share of each dimension, were transformed in barycentric coordinates (x, y),to visualise the results of the AAI, considering the following expressions:

$$x = \alpha + 2*\beta + 1.5*\gamma$$
$$y = \alpha + 1*\beta + (2+\sqrt{3})/2*\gamma$$

with $\alpha+\beta+\gamma=1$.

In geometry, the barycentric coordinate system is a coordinate system in which the location of a point of a simplex is specified as a barycentre of usually unequal masses placed at its vertices.

The objective of the barycentric coordinates is to parameterise a n-simplex using $n+1$ real numbers. In this case, the real numbers are the contribution of teaching (T), research (R) and management (M) to the AAI for each and the n-simplex is an equilateral triangle.

In Figure 2, each point represents a staff member. This graph allows for the identification of similar individuals related to their performance in each one of the dimensions and it helps to interpret global results, which complements the interpretation of each individual value. In this way, this graph is a tool for the interpretation with the same objective of the BoD.

In the university case study, results show four different clusters or profiles obtained depending on the contributions of the three dimensions in the AAI construction of each teacher. Additionally, Figure 3 reports the mean contribution of the three dimensions (in %) in each cluster obtained.

The Management Profile is composed by 2% of the total teachers evaluated (see Table 2). This profile is defined by teachers that have a high mean contribution to the management aspects of the university (80%), but they do also teaching activities (14%) and the 6% of their activity at the university is dedicated to do research activities.

The Teaching-Research Profile is formed by 2.6% of the total teachers evaluated (Table 2) and is composed by teachers with a high mean contribution to the teaching and research activities (55% and 29% respectively) and less but also relevant mean contribution to the management activities (16%). Inside this cluster, we can distinguish between two groups depending on the

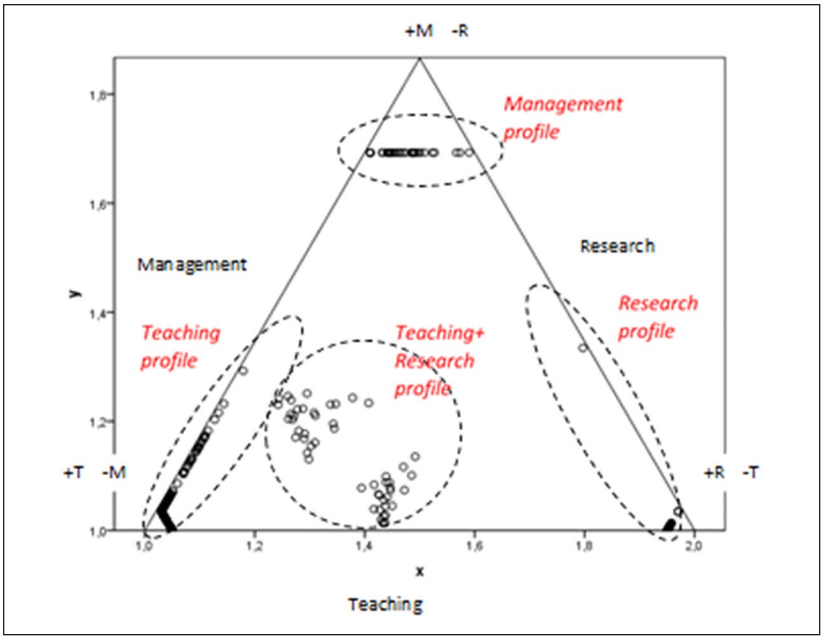

Figure 2. Contributions of the dimensions in barycentric coordinate system.

different intensity in the contribution to teaching and research activities, but we decided to consider both in the same profile.

The Research Profile is composed by 8.1% of teachers (Table 2). Teachers with a clear contribution to the research activities are included in this profile. They teach the minimum ECTS which are bound by the university.

Finally, the Teaching Profile is formed by the majority of teachers evaluated (87,3%). This is in line with the main mission of the university, which correspond to the teaching activities with a mean contribution of 95%.

Table 2 shows the composition of each profile. The categories D, E, F and G are classified in most profiles defined. Teachers that compose these categories have a much more consolidated career than the PhD assistant lecturers. As the results suggest there are teachers within these categories with different teacher profiles. However, the A, B and C categories (which are teachers' categories with a junior academic career) are mainly classified in the Teaching Profile and a low percentage is classified in the Teaching-Research cluster. Furthermore, the Management, Teaching-Research and Research profiles are mainly composed by teachers with an excellent research career (F and G

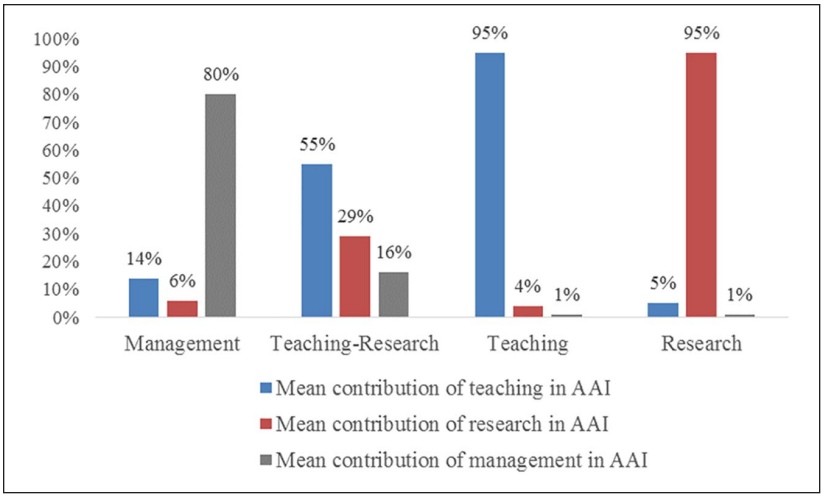

Figure 3. Contribution of the three dimensions in each teacher cluster.

categories). Other studies that research on work environments also suggest that teachers' experiences often differ significantly by rank. For example, newly tenured teachers may find they are required to do more administrative work, thus possibly taking time away from their learning (in doing more research or getting involved in university management tasks) (Neumann & Terosky, 2007). Furthermore, Bozeman and Gaughan (2011) have found that when teachers spend more time thinking and working on their subject matter expertise, they report more satisfaction with their careers and institutions as well as greater research productivity.

As we can see in Figure 4, AAI proposed reaches a maximum value (a value of 1) in the four profiles defined. However, on average, the Research profile has the highest value in AAI. By contrast, the mean of the AAI for teachers classified in the Teaching profile are the lowest one. These results are in line with several studies that found that teachers are likely to be more productive and more satisfied in their careers and with their institutions when they spend more time on research (one part of their learning) (Bozeman and Gaughan, 2011; O'Meara et al., 2017; Winslow, 2010).

Visualising Academic Activity Index

According to Nardo et al. (2005), composite indicator must be able to communicate the final results to decision-makers quickly and precisely. In this

Table 2. Composition of Each Profile According to the Teachers' Category.

		Profile			
Category	%	Management	Teaching-research	Teaching	Research
A	2.2%		[1]2.2%	97.8%	
			[2]1.9%	2.49%	
B	1.2%			95.7%	4.3%
				1.24%	0.6%
C	7.1%			100.0%	
				8.2%	
D	11.8%	0.4%	0.4%	96.6%	2.5%
		2.5%	1.9%	13.1%	3.7%
E	21.2%	0.5%		99.0%	0.5%
		5.0%		24.0%	1.2%
F	41.1%	2.0%	3.4%	88.6%	6.0%
		42.5%	53.8%	41.7%	30.5%
G	15.4%	6.4%	7.0%	52.7%	33.8%
		50.0%	42.3%	9.3%	64.0%
Total	**100%**	2.0%	2.6%	87.3%	8.1%

[1]% teachers in the profile with respect to the total number of teachers in the category.
[2]% teachers in the profile with respect to the total number of teachers in the profile.

study, it is very important to provide graphs to help teachers in the final interpretation of the results. There are different representations to provide information about the results of a composite indicator obtained with BoD methodology (El-Mahgary & Lahdelma, 1995).

In this study, the representation of the results is based on the idea of pie charts. This type of representation was also used for Rogge (2011b) to represent a teaching activity index. The pie chart proposed has the advantage of providing different information in one figure which shows the value of the AAI and the contributions or importance of the teaching, research and management dimensions in the AAI. The value of the overall AAI is presented by the size of the pie, and the contributions of the index by the pie-shares, adding up to the overall score.

The outer circumference drawn in dashed line represents the maximum score that a teacher could obtain in the BoD methodology, a value of 1.

The objective of this pie chart is to provide an idea to a teacher of her/his level of effectiveness in academic activities compared to the benchmark teachers. If the teacher assessed pie chart covers practically the entire outer circumference indicates that the level of relative effectiveness in the

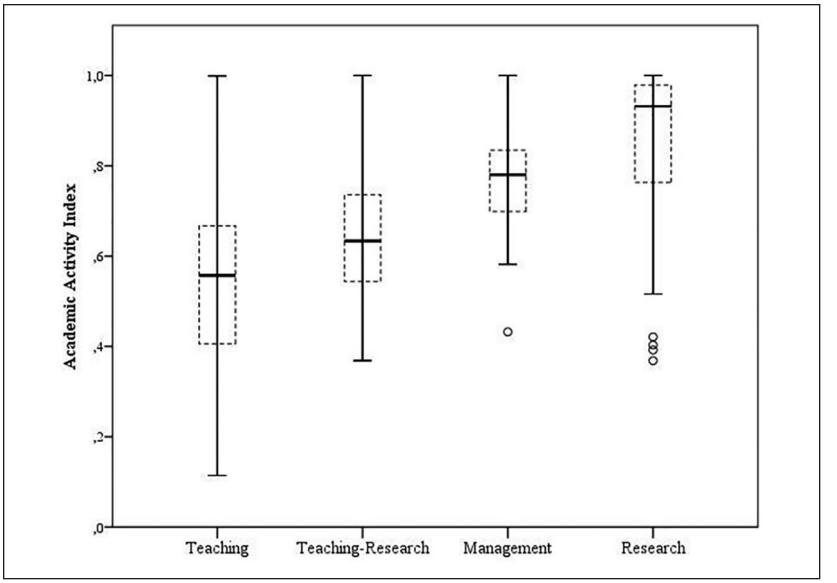

Figure 4. AAI by teacher profiles obtained in barycentric coordinate system.

development of his/her academic activity is like that of the effectiveness of benchmark teachers (Benchmark teachers are those that with the weights of the evaluated teacher obtain an AAI higher than the AAI of the evaluated teacher). By contrast, if the teacher assessed pie chart has a smaller area than the outer circumference, the assessed teacher should improve her/his academic activity to obtain a similar level of effectiveness than the teachers benchmark.

Figure 5 reports, as an example, the visualisation of the AAI of four teachers belonging to different cluster and teachers' categories. For instance, the size of the pie for the represented teacher in (b) (AAI = 0.7455) is larger than the size of the pie for (a) (AAI = 0.6499), but the diverse contribution (pie-shares) of the three main activities without violating the imposed weight restriction illustrate the benefit-of-the-doubt nature of the exercise.

Conclusions and Discussion

The absence of a common methodology, the divergence among teachers on the importance of the simple indicators, the aggregation between indicators, etc., are just ways that reflect the inherent uncertainty in constructing composite indicators.

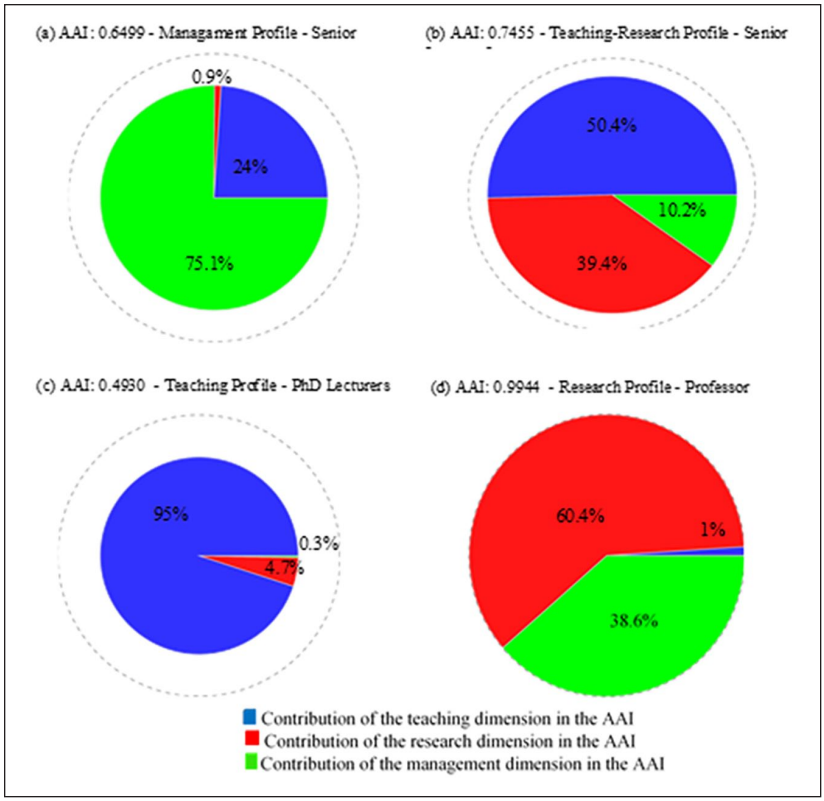

Figure 5. AAI visualization of four teachers.

Several authors found that most reward systems heavily favoured research productivity, even institutions whose primary mission was teaching and/or in institutions that did not have resources to support high levels of scholarly productivity. This conflict of academic functions demoralizes the teaching staff, erodes the vitality of the institution, and cannot help but have a negative impact on students. Furthermore, tenure and promotion committees should consider how to support teachers in developing individualized work plans based on teachers' strengths and orientations (Boyer, 1990; Gonzales and Terosky, 2016).

This study contributes to the literature in this field in applying a Benefit-of-the-Doubt model to develop an Academic Activity Index to measure the academic activity of the university teachers. In contrast to the traditional arithmetic average commonly used in the development of composite indicators, the

Benefit-of-the-Doubt methodology has several benefits when considering the reliability and acceptance of the composite indicators results. First, the BoD-model allows flexibility in the weighting and selects endogenously the weights for the indicators for each using the data set available. Second, BoD-model selects for each teacher the most favourable weights in the evaluation of her/his performance considering her/his teacher profile. Third, BoD-model is flexible to incorporate expert opinion, as we presented in this paper, in the construction of the composite indicator. Fourth, the model is unit invariance, which means that the value of the composite indicator is independent of the unit's measurement of the indicators and, finally, BoD-model allows to compute the contribution or pie-shares of each indicator to the BoD-optimal score providing teachers an idea of the improvement required to achieve the same effectiveness level as observed in the benchmark teachers' performance.

The Academic Activity Index developed using the BoD-model provides the university with a tool that ensure objectivity, and transparency in the definition of the teachers' policies and could be used to develop individualized work plans based on teachers' strengths and orientations considering their different work roles, as suggested by Boyer (1990) and Gonzales and Terosky (2016). Furthermore, the application of the barycentric coordinate system allowed visualising the different profiles of teachers. In this sense, the proposed methodology allows for the design of university policies that help in efficiently managing profiles that differentiate teachers according to their contribution to teaching, research and management activities.

When fixed weights are applied for each one of indicators used in the development of a composite indicator, the evaluation of the different profiles are produced with a common criterion determined because of these fixed weights. If these teachers are evaluated on a regular basis, in the medium term, changes can be made that strengthen the convergence to teacher profiles more balanced, complicating the differentiation. As pointed out by Baker (2017), teacher evaluation systems have become more about promotion and tenure processes rather than being used as developmental opportunities and this could affect the teacher engage in academic work. Although it is logical that each university establishes their own objectives and design policies to enhance certain teacher profiles, it seems interesting to take advantage and develop the dimensions in which each teacher has better performance. It is here, where the assigning of weights proposed in this work can be very helpful since it largely problems that prevents problems created by the assigning of fixed weights.

In the case study in question, four profiles were found: Management Profile, Teaching-Research Profile, Research Profile and Teaching Profile.

In this case, most of the teachers are classified in the teaching profile, which is coherent with the main mission of the university. This information is important and relevant for the university in the establishment of the teacher's policies.

With this proposed method, teachers can be categorized into several profiles, which make it possible to design policies in several lines. For example, in the allocation of teaching load; teachers who have a management profile can reduce the intensity of their teaching activity without their overall assessment being penalized. Similarly, teachers with more teaching prestige and at the same time have less research activity, can be valued precisely for that teaching activity without the low values in their AAI taking away the good work they have done in the teaching part. This can help give value to the teaching activity, which is not always considered with the same degree of importance as researching. However, we must find a balance between this internal way of the university to assess the three activities of a teacher and the requirements of quality agencies in promoting teachers.

The committees that evaluate teachers should have the heterogeneity necessary to be able to evaluate academic activity from the perspective of teaching, research and management. The different actors of the university must also be part of these committees. Therefore, it would be advisable to have professors with proven experience in any of the three activities (teaching, research and management), academic leaders (department directors, deans,. . .), a member of the university's management, students and an expert in evaluation from the pedagogical point of view.

Composite indicators should be accompanied by graphical illustration to facilitate their interpretation. The pie chart representation provides a teacher an idea of her/his level of effectiveness in academic activities compared to the benchmark teachers. Furthermore, the pie chart representation shows the contribution of teachers in each dimension. This information is relevant to teachers so they have a general picture about their academic activity at the university that allows them to establish improvements in the academic activity.

The evaluation of teachers' performance could help universities to provide teachers with opportunities to ensure their professional competence through the design of individualised improvement plans adapted to the needs detected and in terms of the dimension requiring improvement. The results of the study provide an image of the situation of the academic activity of the UPV's teachers, and are available to university managers with the aim to provide a useful tool in the design of policies, motivation for determined activities related to teaching, research and management, and the implementation of improvement plans.

Appendices

Appendix A. Indicators evaluated in the TAI, MAI and RAI index

Table A.1. Indicators Considered in the TAI.

Indicators evaluated in the teaching activity index (TAI)	
Training	Included preparing the faculty for their teaching activities through academic training and educational innovation schemes
Teaching guide	Covered aspects of teaching organisation and educational planning for the subjects taught
Information about tutorials and study groups	Informed the student before the start of the academic year about tutorial options, timetables and the faculty who were teaching the various study groups
Teaching materials	Showed what the lecturer had done to produce teaching materials and resources
Teaching given	Showed what the lecturer had taught
Tutorials carried out	Showed the tutorials the lecturer had given
Assessment	Showed the lecturer's student assessment work
Other teaching activities	Included all other teaching activities a lecturer had did or may do as part of their duties and which were not necessarily related
Student opinion survey	Showed the students' opinion about the lecturer's teaching, especially those aspects related to teaching delivery
Academic performance	Showed the degree to which students had achieved learning objectives in a subject in the context of qualifications, the course and the type of subject
Fulfilment of tutorials	Showed fulfilment of the tutorial schedule by the lecturer during the academic year
Meeting grade report deadlines	Showed compliance with deadlines for the submission of grade reports for subjects

Table A.2. Indicators Considered in the RAI.

Indicators Evaluated In The Research Activity Index (RAI)	
Results of consolidated research	This section evaluated the six-year term and the relevant research awards.
Publication of results in research journals	The number of articles in indexed journals in levels 1,2 and 3 and articles in non-indexed journals were evaluated

(continued)

Table A.2. (continued)

Indicators Evaluated In The Research Activity Index (RAI)	
Publication of results in research congresses	The number of articles or national and international congress presentations were evaluated
Publication of results in Research papers	The number of complete books and the number of chapters in research books were evaluated
Doctoral thesis supervision	The number of doctoral thesis supervised were evaluated
Patents and other registered industrial and intellectual property rights.	The number of patent grants issued internationally or the number of national patents
Artistic creation linked to exhibition spaces	Number of national and international competitive exhibitions
Creative works in different fields	The evaluation of the number of creative works in the field of Architecture and Landscape, the field of Artistic and Literary creation
Achievements of competitive collaborative R+D+i actions	Number of collaborative R+D+i projects and the number of R+D+I under contracts
Raising of financial resources	the number of competitive, collaborative and under contract R+D+i actions; the income generated through Industrial and Intellectual Property Rights License; other actions of technical support under contract; the number of participations in Spin-off of the university were evaluated
General service activities for research	General service activities for research
Other R+D+i merits	Other R+D+i merits

Table A.3. Indicators Considered in the MAI.

Indicators evaluated in the management activity index (MAI)
Participation in the committee of the university community
Coordination of Faculty/Department Directors
Participation in Faculty meeting
Participation in the Department Boards
Participation in the committee of Faculty/Department
Coordination of R+D+i structures

(continued)

Table A.3. (continued)

Indicators evaluated in the management activity index (MAI)
Participation in the committee of R+D+i structures
Participation in the scientific-technical committee for R+D+i structures
Director of Bachelor, Master and / or Doctorate degrees
Participation in the academic committee of Bachelor, Master and / or Doctorate degrees
Participation in the evaluation committee of Bachelor, Master and / or Doctorate degrees
Participation in the claims committee of Bachelor, Master and / or Doctorate degrees
President of the Faculty Board
Participation in Faculty Committee
Participation in professor's negotiation table
Participation in the election Committee of the university
Participation in the Committee for the rules of emeritus Ad Honorem professors
Participation in International Relations committee
Participation in the Committee of valuation of access to study of degree for over 40 years
Participation in the Fair-Trade committee
Participation in the Advisory Board of the Cooperation Centre for Development
Participation in committee for evaluation and follow-up of innovation and convergence projects
Participation in committee of on-line teaching
Participation in the Health and safety Committee

Appendix B. Responsibilities of each teacher's category

Table B.1. Description of the Responsibilities of Each Teacher's Category.

Teacher's category	Description
Non-PhD assistant lecturers (A)	Recruited assistant lecturers with the main purpose of completing their research training. Non-PhD assistant lecturers have a full-time contract. It is a non-tenured position and contracts can range up to four years. They may collaborate in teaching tasks (6 ECTS). The Non-PhD assistant lecturers normally do not taken part in the management activities.

(continued)

Table B.1. (continued)

Teacher's category	Description
Non-PhD lecturers (B)	Lecturers hired by the universities to do exclusively teaching tasks. They have a full-time contract to do teaching (32 ECTS), so they are not required to do research and management activities. The PhD is not a requirement in this category. It is tenured position. The Non-PhD lecturers normally do not taken part in the management activities.
PhD assistant lecturers (C)	Teaching assistants' doctors recruited among junior research doctors with a positive assessment of their activity by the National Agency for Quality Assessment and Accreditation of Spain. They will develop research and teaching tasks (22 ECTS), with full-time contract for up to four years without extension. The PhD assistant lecturers normally do not taken part in the management activities.
PhD lecturers (D)	Teaching doctors that recruited for the development of tasks of teaching (22 ECTS) and research, among doctors that prove at least three years of teaching and research and receive positive evaluation of this activity by the National Agency Quality Assessment and Accreditation. This position requires a research career further consolidated than the PhD assistant lecturers. It is tenured position. The PhD lecturers do not normally taken part in the management activities.
Non-PhD Senior Lecturers (E)	Civil-servant teachers with a full-time for the development of teaching and partial time for researching. They have a contract with an initial teaching capacity of 32 ECTS. The teaching capacity could be reduced by participating in management activities or by giving evidence of the research activities.
Senior Lecturers (F)	Civil-servant teachers with a full-time for the development of teaching (22 ECTS) and research. They have proved the merits and competences for the university teaching body by the National Agency Quality Assessment and Accreditation. They have a senior researching career. They could do management activities that reduce the teaching capacity.
Full Professors (G)	Civil-servant teachers with a full-time contract for the development of teaching (22 ECTS) and research. They are senior lecturer and excellent researcher that have proved the merits and competences for the university teaching body by the National Agency Quality Assessment and Accreditation. In general, they are lecturers that have participated in management activities. Research and university management activities could reduce the teaching capacity in this position. This position is higher in rank to Senior Lecturers.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD


María del Carmen Bas https://orcid.org/0000-0002-1221-0641

References

Allen, R., Athanassopoulos, A., Dyson, R. G., & Thanassoulis, E. (1997). Weights restrictions and value judgements in Data Envelopment Analysis: Evolution, development and future directions. *Annals of Operations Research*, *73*(0), 13–34.

Arneson, R. J. (1990). Primary goods reconsidered. *Noûs*, *24*(3), 429–454.

Avital, M., & Collopy, F. (2001). Assessing research performance: Implications for selection and motivation. *Sprouts: Working Papers on Information Environments*, *1*(3): 40–61.

Baker, V. L. (2017). Faculty members' scholarly learning across institutional types. *ASHE Higher Education Report*, *43*(2), 9–138.

Bana e Costa, C., & Oliveira, M. (2011). A multicriteria decision analysis model for faculty evaluation. *Omega - The International Journal of Management Science*, *40*(4), 424–436.

Bas, M. C. (2014). Estrategias metodológicas para la construcción de indicadores compuestos en la gestión universitaria (Doctoral dissertation). Faculty of Statistics and Operational Research Applied and Quality, Universitat Politècnica de València.

Bas, M. C., Tarantola, S., Carot, J. M., & Conchado, A. (2017). Sensitivity analysis: A necessary ingredient for measuring the quality of a teaching activity index. *Social Indicators Research*, *131*(3), 931–946.

Beasley, J. E. (1995). Determining teaching and research efficiencies. *Journal of the Operational Research Society*, *46*(4), 441–452.

Blackburn, R. T., & Lawrence, J. H. (1995). *Faculty at work: Motivation, expectation, satisfaction*. The Johns Hopkins University Press.

Boyer, E. L. (1990). *Scholarship reconsidered: Priorities of the professoriate*. Carnegie Foundation for the Advancement of Teaching.

Bozeman, B., & Gaughan, M. (2011). Job satisfaction among university faculty: Individual, work, and institutional determinants. *The Journal of Higher Education*, *82*(2), 154–186.

Campbell, C. M., & O'Meara, K. (2013). Faculty agency: Departmental contexts that matter in faculty careers. *Research in Higher Education*, *55*(1), 49–74.

Cano, J. J., Carot, J. M., Fernández-Prada, M. A., & Fargueta, F. (2009). An evaluation model of the teaching activity of academic staff. In *Proceedings of I conference quality of teaching in higher education*. Turkey.

Charnes, A., Cooper, W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, *2*(6), 429–444.

Cherchye, L., Moesen, W., Rogge, N., & Van Puyenbroeck, T. (2007). An introduction to 'Benefit of the Doubt' composite indicators. *Social Indicators Research*, *82*(1), 111–145.

Cherchye, L., Moesen, W., Rogge, N., Van Puyenbroeck, T., Saisana, M., Saltelli, A., Liska, R., & Tarantola, S. (2008). Creating composite indicators with DEA and robustness analysis: the case of the Technology Achievement Index. *Journal of the Operational Research Society*, *59*, 239–251.

De Witte, K., & Rogge, N. (2011). Accounting for exogenous influence in performance evaluations of teachers. *Economics of Education Review*, *30*(4), 641–653.

De Witte, K., Rogge, N., Cherchye, L., & Puyenbroeck, T. (2013). Economies of scope in research and teaching: A non-parametric investigation. *Omega*, *41*(2), 305–314.

El-Mahgary, S., & Lahdelma, R. (1995). Data envelopment analysis: Visualizing the results. *European Journal of Operational Research*, *83*(3), 700–710.

Emery, C. R., Kramer, T. R., & Tian, R. G. (2003). Return to academic standards: A critique of student evaluation of teaching effectiveness. *Quality Assurance in Education*, *11*(1), 37–46.

Fleurbaey, M. (1992). Social choice and the indexing dilemma. *Social Choice Welfare*, *29*(4), 633–648.

Giménez, V., & Martínez, J. L. (2006). Cost efficiency in the university: A departmental evaluation model. *Economics of Education Review*, *25*(5), 543–553.

Gonzales, L. D., & Terosky, A. L. (2016). From the faculty perspective: Defining, earning, and maintaining legitimacy across academia. *Teachers College Record*, *118*(7), 1–44.

Johnes, G., & Johnes, J. (2009). Higher education institutions'cost and efficiency: Taking the decomposition a further step. *Economics of Education Review*, *28*(1), 107–113.

Marsh, H.W., & Roche, L. (1994). *The use of students' evaluation of university teaching to improve teaching effectiveness*. Australian Department of Employment, Education, and Training.

Melyn, W., & Moesen, W. (1991). *Towards a synthetic indicator of macroeconomic performance: unequal weighting when limited information is available* (Public Economic Research Paper 17). Center for Economic Studies.

Nardo, M., Saisana, M., Saltelli, A., & Tarantola, S. (2005). *Tools for composite indicators building* (European Commission, report EUR 21682 EN (EUR collection)). Joint Research Centre.

Neumann, A., & Terosky, A. L. (2007). To give and to receive: Recently tenured professors' experiences of service in major research universities. *The Journal of Higher Education*, *78*(3), 282–310.

O'Meara, K., Rivera, M., Kuvaeva, A., & Corrigan, K. (2017). Faculty learning matters: organizational conditions and contexts that shape faculty learning. *Innovative Higher Education*, *42*(4), 355–376.

O'Meara, K., Terosky, A.L., & Neumann, A. (2008). Faculty careers and work lives: A professional growth perspective. *ASHE Higher Education Report*, *34*(3), 1–221.

Pina, V., & Torres, L. (1995). Evaluación del rendimiento de los departamentos de contabilidad de las universidades españolas. *Hacienda Pública Española*, *135*(-), 183–190.

Pritchard, R. D., Watson, M. D., Kelly, K., & Paquin, A. R. (1998). *Helping teachers teach well: A new system for measuring and improving teaching effectiveness in higher education*. The New Lexington Press.

Ramón, N., Ruiz, J. L., & Sirvent, I. (2010). Using Data Envelopment Analysis to assess effectiveness of the process at the university with performance indicators of quality. *International Journal of Operations and Quantitative Management*, *16*(4), 87–103.

Rawls, J. (1999). *A theory of justice*. Oxford University Press.

Rogge, N. (2011a) Granting teachers the "benefit of the doubt" in performance evaluations. *International Journal of Educational Management*, *25*, 590–614.

Rogge, N. (2011b). *Evaluating teaching and research performances: a non-parametric 'Benefit of the Doubt' approach* (Doctoral dissertation). Faculty of Bussiness and Economics, Katholieke Universiteit Leuven.

Roll, Y., Cook, W. D., & Golany, B. (1991). Controlling factors weights in Data Envelopment Analysis. *IIE Transactions*, *23*(1), 2–9.

Roll, Y., & Golany, B. (1993). Alternate methods of treating factor weights in DEA. *Omega*, *21*(1), 99–109.

Saltelli, A. (2007) Composite Indicators between analysis and advocacy. *Social Indicators Research*, *81*, 65–77.

Sharpe, A. (2004). Literature review of frameworks for macro-indicators (CSLS research report 2004-03). Centre for the Study of Living Standards.

Schuster, J. H., & Finkelstein, M. J. (2006). *The American faculty: The restructuring of academic work and careers*. Johns Hopkins University Press.

Thanassoulis, E., & Allen, R. (1998). Simulating weights restrictions in Data Envelopment Analysis by means of unobserved DMUs. *Management Science*, *44*(4), 586–594.

Thanassoulis, E., Portela, M. C., & Allen, R. (2004). Incorporating value judgments in DEA. In W. Cooper, L. Seiford, & J. Zhu (Eds.), Handbook on data envelopment analysis (pp. 99–138). Kluwer Academic Publishers.

Thompson, R., Langemeier, L., Lee, C.-T., Lee, E., & Thrall, R. (1990). The role of multiplier bounds in efficiency analysis with application to Kansas farming. *Journal of Econometrics*, *46*(1–2), 93–108.

Tomkins, C., & Green, R. (1988). An experiment in the use of data envelopment analysis for evaluating the efficiency of UK university departments of accounting. *Financial Accountability and Management*, *4*(2), 147–164.

Vicerrectorado de Calidad y Evaluación de la Actividad Académica (2014). Manual de evaluación de la actividad docente del profesorado de la Universidad Politécnica de Valencia. Retrieved from http://www.upv.es/entidades/VECA/menu_urlc.html?/entidades/VECA/info/U0668892.pdf

Vicerrectorado de Investigación, Innovación y Transferencia (2008). Reglamento para la evaluacion de la actividad de investigación, desarrollo e innovación en la Universidad Politécnica de Valencia. Retrieved from http://www.upv.es/entidades/VIIT/menu_urlc.html?/entidades/VIIT/info/U0451850.pdf

Winslow, S. (2010). Gender inequality and time allocations among academic faculty. *Gender and Society*, *24*(6), 769–793.

Wong, Y.-H. B., & Beasley, J. E. (1990). Restricting weight flexibility in data envelopment analysis. *Journal of the Operational Research Society*, *41*(9): 829–835.

Author Biographies

María del Carmen Bas is working as a teacher in the Business Mathematics Department at the Universitat de València. She obtained her Mathematical Science degree and her Ph.D. in Statistics and Optimization. Her main researching areas are the multivariate analysis, the design and analysis of surveys, the sample design, the generation of data bases and the application of sensitivity and uncertainty analysis to the construction of composite indicators.

José Miguel Carot studied mathematic science, specializing in Statistics and Operations Research, at the Universidad de Valencia, and obtained his PhD from the Universidad Politécnica de Valencia (UPV). He is Professor at UPV in the Department of Applied Statistics and Operational Research and Quality. He researches and works in higher education management, survey methodology and multivariate methods, participating in and directing 28 public research projects.