ORIGINAL

QUESTIONNAIRE ON METHODOLOGY AND ASSESSMENT IN PHYSICAL EDUCATION INITIAL TRAINING

CUESTIONARIO SOBRE METODOLOGIA Y EVALUACIÓN EN FORMACIÓN INICIAL EN EDUCACIÓN FÍSICA

Castejón Oliva, F.J.[1]; Santos Pastor, M.L.[2] and Palacios Picos, A.[3]

[1] Universidad Autónoma de Madrid. Department of Physical Education, Sport and Human Movement. Tenured Professor, javier.castejon@uam.es
[2] Universidad Autónoma de Madrid. Department of Physical Education, Sport and Human Movement. Tenured Professor, marisa.santos@uam.es
[3] Universidad de Valladolid. Facultad de Educación de Segovia. Tenured Professor, palacios@psi.uva.es

Spanish-English translators: SG Linguistic Consulting S.L. http://sglinguistics.com.

ABSTRACT

This article describes the validation process of a scale aimed at knowing the perceptions regarding participative methodologies and the formative evaluation of Physical Education university students. We have undertaken a process of validation of the contents, comprehensibility and reliability of the questionnaire in three stages: In the first one, expert validation with six expert judges in university evaluation; in the second, comprehensibility validation with a sample of 50 students; in the third, reliability and general validation, with 892 university students from 10 Spanish universities. Results show the reliability of the instrument and the confirmation with the factorial analysis. The scale has turned out to be a useful instrument to help teachers analyse students' perceptions regarding participative methodologies and

formative evaluation. We hope it becomes a useful tool to redirect modifications and improvements in university teaching.

KEYWORDS: Questionnaire, Validation, Formative assessment, Participative methodology, Initial training, Physical education, Higher education.

RESUMEN

Este artículo describe el proceso de validación de una escala para conocer las percepciones sobre metodologías participativas y de evaluación formativa de estudiantes universitarios de titulaciones de Educación Física. Se ha realizado un proceso de validación de contenido, comprensión y fiabilidad del cuestionario en tres fases: En la primera, de validación de expertos con seis jueces expertos en evaluación universitaria; en la segunda, de validación de la comprensión, con una muestra de 50 estudiantes; en la tercera, de fiabilidad y validación general, con 892 estudiantes universitarios de 10 universidades españolas. Los resultados han mostrado la fiabilidad del instrumento y la confirmación con el análisis factorial. La escala resulta ser un instrumento válido para ayudar al profesorado a analizar las percepciones de los estudiantes en relación con las metodologías participativas y la evaluación formativa. Se espera que sea una herramienta eficaz para orientar el cambio y la mejora en la docencia universitaria.

PALABRAS CLAVE:: Cuestionario, Validación, Evaluación formativa, Metodología participativa, Formación inicial, Educación física, Educación superior.

INTRODUCTION

These last few years have seen a clear interest in generating changes in university teaching regarding both its objectives and methodological guidelines, as well as in the assessment of teaching-learning processes, as announced in the objectives of the European Council of Lisbon 2000 (European Council, 2000; EEES, 2009). An improvement in methodological implementation and in the application of differentiated assessment methods is called for (Vázquez, 2008) in order to achieve university excellence (Villa, 2008), an excellence that must be put into question when "around 80% of the assessments performed throughout the world is done in the shape of exams and essays" (Brown and Glasner, 2007, p.8).

With the development of university degrees after the approval of the new curricula, there is a clear proposal for changes regarding teaching in university. Amongst the most significant aspects of this changes we find methodology (Biggs, 2006; Tejedor, 2003), which, guided by the concept of the search for university excellence, implies: (a) proposing a participative methodology, linked to an active learning (Millis, 2010; Meyer and Jones, 1993); (b) promoting a communication between teachers and students that includes reflexive processes to know, understand, analyse, apply, synthesize and evaluate (Black and Wiliam, 2009; Huber, 2008); (c) establishing links between theory and practice, and integrating knowledge in order to provide meaning to what has been learnt (Cano-González, 2009; Huber, 2008); (d) searching for different methods of organising students and their tasks (Marín-García, Miralles-

Insa, García-Sabater and Vidal-Carreras, 2008); and (e) using information technology instruments and media (Cano-González, 2009).

In this new framework, the evaluation must be linked to the concept of formative assessment, which will allow us to verify the results of the learning process, and not only the results that consider assessment a synonym for grades; it will also become an element of improvement and a help for teachers and students (MacMillan, 2007; Pérez, Julián and López-Pastor, 2009).

Studies show that evaluation and students' perceptions of the teaching-learning process are after all a way for teachers to improve their teaching (Marsh, 1987, 2007; Marsh and Roche, 1997), and it seems necessary to study them in depth as a way of ensuring the quality of teaching (Villa, 2008). The way in which students understand their own learning is conditioned by their own beliefs about how people learn (Tippin, Lafreniere and Page, 2012), but it is also a result of how they themselves acquired the knowledge with the experience of their own training (Irons, 2008).

However, there are few studies devoted to the perception of students of evaluation processes. We may underline here the contributions of Marsh (1982) and the *Students' E+valuations of Educational Quality* (SEEQ), a study that considers eight aspects related to teaching methodologies and formative assessment: (a) relationship between learning and its value; (b) eagerness of the teacher; (c) group interaction; (d) individual reports; (e) relationship between organization, clarity and scope of proofs; (f) relationship between exams and annotations; (g) readings and work documents; and (h) relationship between work load and difficulty.

Previous studies on methodology and teaching in university were centred on so-called effective teaching, based on the following key elements: the interest and motivation of teachers, respect toward the students' learning pace, the adoption of clear objectives, the control of the students' progress and being able to learn from students themselves (Race, 2007); not forgetting other factors centred on the teachers' personalities (Feldman, 1986; Murray, Rushton and Paunnem, 1990).

Another study based on the learning evaluation systems that teachers use is *Shortened Experiences of Teaching and Learning Questionnaire* (SETLQ) (University of Edinburgh, 2001). Biggs (1987) uses as foundation the learning style of students and the context of the teaching activity; the work of Entwistle and collaborators (Entwistle, Hanley and Hounsell, 1979; Entwistle and Peterson, 2004; Entwistle, Tait and McCune, 2000) with *Approaches and Study Skills Inventory for Students* (originally designed by Ramsden and Entwistle, 1981) shows differences regarding how studying and learning are conceived (Cano-García and Justicia-Justicia, 1994). They all emphasize as a shared idea a clear relationship between the context, the proposed activity, the expected results and the students' predisposition (Bain, 2005; Cabrera and La Nasa, 2002).

In any case, the interest of teachers to know their very own practice implies paying attention, among other things, to how teaching and assessing are conceived, without need for institutional controls (Fuentes-Medina and Herrero, 1999; Stenhouse, 1987; Knight, 2005; Leathwood and Phillips, 2000; Ramos and Ortiz, 2010).

In the specific case of the teachers' initial training a greater attention has been called for regarding the methodologies and evaluation systems used in university teaching, precisely because future teachers are being taught (Gimeno, 2012).

However, studies on this subject are scarce. Even in those in which questionnaires on participative methodologies and formative evaluation have been used, they have been used more as a means than an end, and very few resources have been devoted to their psychometric characteristics. Palacios and López-Pastor (2013) have developed a Scale of Attitudes towards Formative Evaluation and the Participation of Students (EAEF-PA in Spanish) aimed at measuring the attitudes of students towards the use teachers make of formative evaluation. Also, Gutiérrez-García, Pérez-Pueyo, Pérez-Gutiérrez and Palacios-Picos (2011) have prepared a questionnaire related to the formative evaluation of students in which they check students' participation in the evaluation, their perception of the use of methodological strategies and techniques, the progress of classes and the evaluation of the professional competences of teachers.

However, and as we have already pointed out, in spite of the importance of having information about the possible change toward more rich methodological systems and more formative assessment systems, as of today there are no valid and reliable query instruments regarding these changes, at least in opinion of one of the key elements of those processes, students.

OBJECTIVE

Our goal has been to design and verify a scale of perception regarding participative methodologies and formative evaluation (EMPEF in Spanish) which will allow us to assess the way in which methodology and evaluation are perceived by students during their initial training, as well as the degree of satisfaction achieved in their implementation in the learning process.

METHOD

Design

We have undertaken a process of validation of the contents, comprehensibility and reliability of the questionnaire in three stages: A first stage done with experts in the field, a second stage in which a pilot study is done with the aim of specifying the questionnaire, and a third stage in which the questionnaire is applied to a majority group of students of several universities (Figure I).
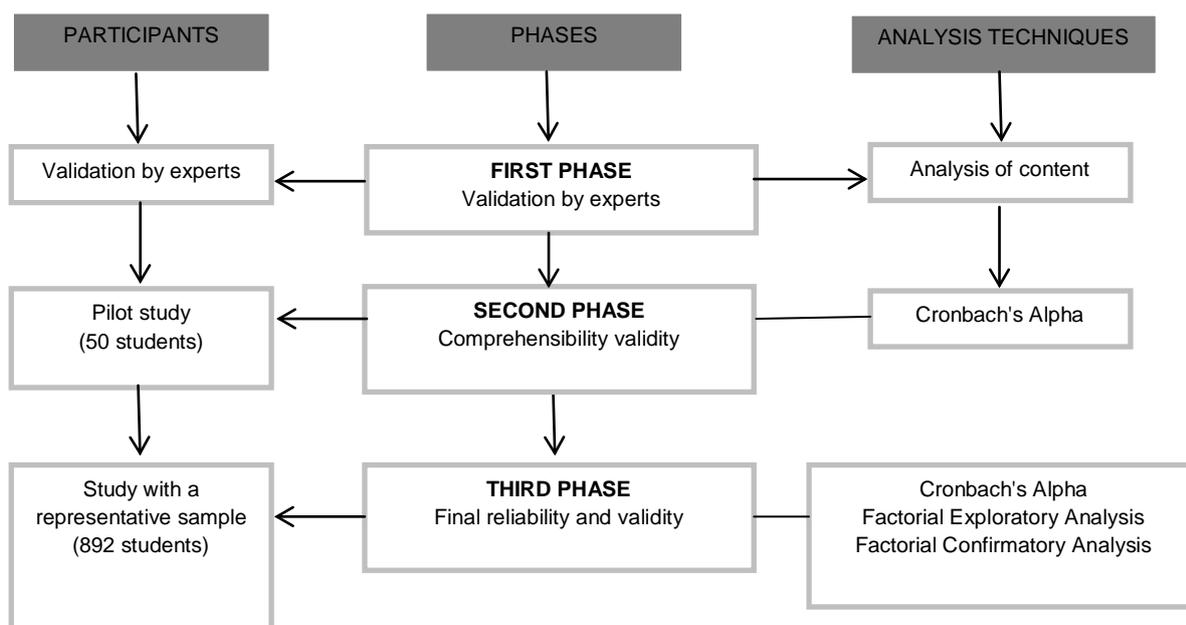
Figure I.- Stages of the process, participants and statistical techniques used to elaborate the EMPEF.

Participants

In the first stage six expert judges in university evaluation took part. They have been teaching for more than 10 years in public universities, in the Specialist Teacher in PE degree and the Degree on Physical Activity and Sport Sciences. They all have PhD's and have published articles and books on formative evaluation and participative methodologies in university teaching. It was an intentional selection.

In the second stage, aimed at obtaining the validity of comprehensibility and a first advancement in the reliability of the instrument, we have selected intentionally 50 university students on several courses of the two most typical degrees in PE: PE Specialist Teacher and the Degree on Physical Activity and Sport Sciences. The selection of the sample was intentional.

The implementation of the third stage was done with a sample of 892 university students of 10 Spanish universities, taking courses on Teacher degrees and the Degree on Physical Activity and Sport Sciences. The sample included all courses in each degree.

Procedure

 The questionnaire was given to the group of experts in order for them to verify the degree of coincidence/divergence of the terms used and the assessment of the questions in all sections, which helped configure a first version of the questionnaire. This first version was handed to a group of 50 students, representative of the final population of the study, which helped to check the comprehensibility of the terms and the validity of the field, as well as its reliability and internal consistency. In a third stage, the final version of the questionnaire was applied to the complete sample of

students, with the corresponding permissions and keeping their anonymity at all times.

DATA ANALYSIS

With the final data matrix, the participative methodology and formative evaluation questionnaire factors in university teaching were extracted by means of a Factorial Exploratory Analysis (AFE) using the SPSS 18.0 statistical software. The method for extracting factors used was principal components with orthogonal rotation (Varimax). Also, the reliability of the scale was reviewed by means of Cronbach's Alpha statistic. In order to verify in a more precise manner the validity a Factorial Confirmatory Analysis (AFC) was made by means of the Lisrel 8.6 software program.

RESULTS

First stage: Content validity

In order to verify the content validity of the scale, the questions were established following relevance and representativeness criteria regarding the population targeted and its specific domain (Latiesa, 2000; Losada and López-Feal, 2003) and taking into account previous studies related to the subject (Fernández-Pérez, 1989; Gutiérrez-García et al., 2011; Palacios and López-Pastor, 2013; Prieto, 2007; Tejedor, 1998; Trillo, 2005; Trillo and Porto, 1999). After the pertinent adjustments, the initial scale was given to the experts, following a previous consent of their participation and their authorization. In a small presentation they were informed of the goal of the questionnaire. The experts assessed their degree of agreement with each of the questions of the scale with a rate between 1 and 4 (3 and 4 agreement, 1 and 2 disagreement). In case of disagreement, an alternative answer was requested. If there were no alternatives, the question was removed from the scale. In this manner a degree of exactness and correction was obtained for the wording of each of the items.

The biggest disagreement happened when differentiating between *methodology* and *evaluation* (following the proposal of the "aligned teaching" concept of Biggs, 2006). For instance, one of the expert evaluators (their anonymity is kept by means of codes) declared:

> Do students know what it is they are being asked? Do they know the meanings of all these concepts? However, in this section of the evaluation I miss other important concepts: 'they –the teachers– return corrected essays in a short span of time'; 'it is allowed to rewrite an essay or document and hand it in again', etc. (EvExp1).

There were also doubts regarding the wording of some questions related to the scope of the evaluation (in one or in several subjects):

> The questionnaire is meant to assess the subject matter, not the plan or the centre. Although it might be interesting to bring this up, to verify the level of internal coordination (it might be interesting), then the text would have to

change slightly: "There are similarities with the evaluation systems of other subject matters in this same centre" (EvExp3).

On the other hand, our intention was to use a Likert-like scale of 5 alternatives (1 to 5), but all experts suggested using values from 0 to 4, "as perception and measurements are more precise, as well as the equivalent to nothing-none, better with 0 than with 1" (EvExp4).

Second stage: Comprehensibility validity

With the modifications performed in the previous section, the test was used on two groups of 50 students, half studying the Specialist Teacher in Physical Education Degree and the other half the Degree on Physical Activity and Sport Sciences. In all cases they were asked for permission to participate in this study. The objective at this stage was to verify whether the type of questions being asked was the most adequate and the degree of comprehension of the wording. We also intended to check the total duration of the scale, and whether it was inside the acceptable limits for this kind of tests.

The duty of the students was to assign a score to the degree of comprehension they had regarding the offered items. In order to attain a qualitative comprehensive degree of the questionnaire, they were asked to write down doubts, questions and suggestions. They were asked, finally, to point out which items they did not understand. Those responses written down with a No opinion/No reply by more than 5% of the sample were directly removed. On the other hand, those questions that offered values with the same response over 90% were removed as well. In this way we managed to use those questions that made a difference regarding a high or low assessment.

An analysis of the reliability was also made, in order to verify the consistency of the measurement when it was repeated in several occasions. For the calculation of the reliability Cronbach's Alpha coefficient was used. The results obtained in this second stage are pretty high. For the whole scale, the reliability index was 0.92.

For the calculation of reliability in this stage questions 1 (with three secondary questions, one about negotiation and two about competences), 14, 15 and 16 were not included, as they were questions independent of the construct to be measured (perception of participative methodologies and formative evaluation) but contributed with valuable information. Their wording and the results obtained in this second stage are summarized in table I.

Questions	Percentages				
	Not at all	A little	Quite a bit	A lot	No opinion/No reply
Item 1a. Was the subject matter's program negotiated at the beginning of the course?	10.4	25.0	27.1	16.7	20.8
Item 1b. Has the methodology used in this subject helped you acquire professional competences?	6.6	20.6	40.0	32.2	0.6
Item 1c. Does the evaluation outlined favour the acquisition of professional competences?	7.2	21.5	45.3	24.4	1.6
Item 14. Indicates the overall satisfaction regarding the subject	6.4	25.5	36.2	19.1	12.8
Item 15. Indicates the overall satisfaction regarding the subject's evaluation	2.1	10.4	31.3	29.2	27.0
Item 16. Indicates the degree of difficulty of the subject	8.6	32.2	39.8	3.8	15.6

Table I. Distribution of responses to questions not related directly to the construct represented by the questionnaire.

Item 1 reflects the negotiation of the program, and its presence is deemed important due to the need to verify whether there is a negotiation between teacher and student. Items 14, 15 and 16 reflect global valuations of the subject that might be taken into account in case a calculation could be made of the correlations between evaluation and methodology with the degree of difficulty expressed. At the end of the questionnaire there is a question open so that students may make comments or observations.

Third stage: Final reliability and validity

Once the questionnaire was closed as can be deduced from the pilot study of the second stage, it was filled by a wide group of students with the intention of verifying its validity and reliability. As we have previously mentioned, in this third stage 892 students from 10 universities took place, studying Teaching degrees or the Degree on Physical Activities and Sports Sciences. In all cases, permission was asked and granted to participate in the study (to teachers, university and students alike).

The group is distributed with a greater percentage of men (55.1%) than women (44.9) and an average age of 21 years and 6 months. The majority of students that responded to the questionnaires had a high degree of attendance, and that is why we may conclude that the information they provided suggests a knowledge of what is done daily in classes, both regarding methodology and evaluation.

Scale of participative methodologies. A new calculation was then done with the whole sample of the reliability indexes of the Scale of participative methodologies using as foundation Cronbach's Alpha. As we may see in table II, the 0.84 value obtained allows to ensure measurements of high reliability. We decided to keep all of the scale's original items, as removing any of them did not result in a significant increase in the value of final reliability.

Cronbach's Alpha	Cronbach's Alpha based on the typified elements	Number of elements
.838	.845	24

Table II. Reliability index of the Scale of participative methodologies

Once the scale's reliability was ensured, the next step was to analyse its implicit factors by means of a Principal Components Factorial Analysis. The results of this analysis are summarized on table III. Adequate values are obtained both in the KMO index, 0.807, and in Barlett's sphericity test (p>.00).

	Component							Factors
	1	2	3	4	5	6	7	
03. Books or articles reviews	.683							F1 Methodologies based on written reports and journals
04. Reports	.673							
02. Journal or session sheets	.593							
01. Field notebooks	.574							
09. Checking bibliography	.521							
06. Notebook or similar to write down notes	.511							
02. Questions were allowed in classes		.806						F2 Methods that strengthen human relations
01. A positive classroom climate was favoured		.763						
03. Tutorials were used to solve doubts about the subject's progress		.587						
04. Audiovisual media have been used in classes (transparencies, presentations, videos, etc.)			.796					F3 Classical methodologies with audiovisual and ICT support
05. Virtual platforms			.657					
07. Oral expositions were used on classes to give explanations, and notes were taken			.564					
06. Guest speech			.526					
02. Seminar				.834				F4 Exhibition techniques
03. Case studies				.722				
11. Round tables				.595				
05. The subject was accompanied by practices that helped to understand better the future professional work					.740			F5 Use of internships
10. Observation of teaching practices					.685			
06. A variety of methodological strategies was used					.495			
07. Dialogic meetings						.7		F6

253

08. Debates							92.754	Discursive methodologies
01. Tutelary projects							.666	F7 Group methodological techniques
05. Collaborative learning							.591	
04. Learning based on problems / problem solving							.410	
Auto-values	12.15	3.72	2.03	1.58	1.23	1.15	1.01	
% Explained variance	31.98	9.79	5.36	4.17	3.24	3.02	2.64	
% Accumulated explained variance	31.98	41.77	47.14	51.32	54.56	57.59	60.23	
KMO: .807								
Barlett' sphericity test: Chi-approximate square: 3808.87. gl. 276; next: .000								

Table III. Factorial Analysis of the Scale of participative methodologies

The first of the factors in the questionnaire (F1) is the one that accounts for the greater amount of variance (32%); it bears great significance in five questions related to the preparation of reviews and reports, with bibliographic queries and with the elaboration of journals and field notebooks; for all these reasons, we have called this factor *Methodologies based on written reports and journals*.

The second factor, which would account for 10% of the scale's total variance, has an important interpersonal component with high factorial effects in the following questions: "Questions were allowed in classes", "A positive classroom climate was favoured" and "Tutorials were used to solve doubts about the subject's progress". We have called this factor *Methodologies that strengthen interpersonal relations*.

The third factor (F3) has effects in features such as the use of audiovisual media in classes and guest speeches. For that reason we have called this factor *Classic methodologies with audiovisual and ICT support;* it accounts for 4% of the total variance.

We have called the fourth factor (F4) *Innovative exhibition methodological techniques and case studies* due to their significant values in questions related to the use of these study strategies, seminars and round tables.

Factors 5 and 6 have values in only two items each, with an accounted variance of 3% each. The first of them has a clear practical component, as methodological strategies with significant values for the questions: "The subject was accompanied by practices that helped to understand better the future professional work" and in the "Observation of teaching practices". The second of them, for its part (F6), has clear high effects on the use of "dialogic meetings" and "debates", and for that reason we have called it *Discursive methodologies*.

This last factor, with an autovalue over 1, accounts for 3% of the total variance and is correlated in a significant manner with three items, specifically: "Tutelary projects",

"Collaborative learning" and "Learning based on problem solving". We have called it *Group methodological techniques.*

In order to ensure the validity of the content, we performed a Factorial Confirmatory Analysis (AFC) of the seven factors already described. The indexes obtained from the covariance matrix had satisfactory settings: both in the RMSEA = 0.078 index and in the GFI = 0.88 index, and also in the remaining values to set up the model (Table IV).

Model	S-B$_{(Chi-square)}$ (gl) (p)	GFI	RMSEA	NFI	NNFI	CFI	AGFI	AIC
Scale factors (seven factors)	1358.91 (231)(P = .00)	.88	.078	.86	.86	.88	.84	10194.78

Table IV. Evaluation of the seven-factor model of Participative methodologies scales by means of AFC

Evaluation systems scale*.* Just as we did with the Participative methodologies scale, we have performed an analysis of the reliability of the Evaluation systems scale by means of Cronbach's Alpha, results which we summarize on table V.

Cronbach's Alpha	Cronbach's Alpha based on the typified elements	Number of elements
.839	.868	38

Table V. Reliability index of the Evaluation systems scale

The 0.83 Alpha obtained allows us to conclude that this is a scale with a very high reliability level. Just like it happened with the previous scale, we decided not to remove any of the 38 items of the original scale, as that removal did not imply values significantly greater in the total scale.

The Principal Components Factorial Analysis (AFCP) obtained six factors that would account for 60% of the scale's total variance. Both the .087 KMO index and Barlett's sphericity test (p>.00) indicate the importance of this analysis and the statistical significance of its results (Table VI).

	Component						Factors
	1	2	3	4	5	6	
11. Much more is learnt(*)	.799						F1
09. Allows functional learnings	.779						Evaluation oriented to learning processes
10. Generates significant learnings	.776						
12. Improves the quality of requested essays	.774						
08. Improves academic tutelage (follow-up and help for students)	.765						
06. The student is more motivated, and the learning process is more motivational	.746						
07. Grades are more fair	.743						
14. Evaluates all possible aspects	.733						

Item							Factor
13. There is a correlation between theory and practice	.673						
15. There is feedback and the possibility to correct mistakes in essays and activities	.650						
04. The student performs an active learning	.649						
16. A more personalized follow-up is provided	.629						
01. It offers alternatives for all students	.628						
17. Requires more responsibility	.611						
03. It is centred on the process, the importance of daily work	.575						
02. There is a previous contract, agreed and negotiated, regarding the evaluation system	.559						
05. Teamwork is conceived in a collaborative manner	.504						
11. It is unfair compared to other evaluation processes		.713					F2 Anti-evaluation
10. Generates uncertainty and insecurity, doubts about what is to be done		.698					
12. Corrections were not clear enough		.670					
09. The process is more complex and sometimes unclear		.664					
13. The assessment of the work done is subjective		.592					
02. Formative evaluation processes were used (the teacher corrected activities or documents, and provided feedback as to how to improve and correct mistakes)			.748				F3 Formative evaluation
03. The observations made regarding essays or required activities might result in their repetition in order to improve them			.745				
01. Continuous evaluation processes were performed (activities or documents assessments during the subject's duration)			.665				
07. A lot of work may be accumulated towards the end				.751			F4 Non-planned evaluation
08. The relationship work/credits is disproportionate				.640			
05. Demands a greater effort				.574			
06. It is difficult to work in teams				.548			
01. Demands compulsory and active attendance					.817		F5 Attendance-based evaluation
03. Demands continuity					.715		
04. A portfolio or individual folder was used to hand in materials						.826	F6 Portfolio-based evaluation
05. A group portfolio or collaborative folder was used to hand in materials						.748	
07. There were similarities with the evaluation systems of other subject matters in this same centre						.560	
Auto-values	12.1	3.72	2.04	1.58	1.23	1.15	

	5					
% Explained variance	31.98	9.79	5.36	4.17	3.24	3.02
% Accumulated explained variance	31.98	41.77	47.14	51.32	54.56	57.59
KMO: .807						
Barlett' sphericity test: Chi-approximate square: 3808.87. gl. 276; next: .000						

Table VI. Factorial Analysis of the Evaluation systems scale (The 4 items with factorial values under 0.35 do not appear on the list)

The first factor, the most numerous with factorial weightings in a total of 19 questions, accounts for 32% of the total variance. In spite of the heterogeneity that these questions show (evaluation methods that favour learning and significant learning, the improvement of interpersonal and didactic relations, the use of formative evaluation processes, etc.), they all have in common that they might perfectly be indicators of an *Evaluation oriented to learning processes*.

The second factor is where the highest correlations were obtained in the questions: the evaluation system employed is unfair compared to other evaluation processes, it generates insecurity and uncertainty, with not very clear corrections and equally confusing processes and a high degree of subjectivity in assessments. For all these reasons we have called it *Anti-evaluation*.

The third factor accounts for 5% of the total variance of the questionnaire with high correlations for the following questions: formative evaluation processes were used, the observations made regarding essays or required activities might result in their repetition in order to improve them; this factor has been identified as *Formative Evaluation.*

The fourth factor found was related to four of the questionnaire questions dealing with an inadequate temporal organization of the evaluation processes, with an elevated perception of resource assignment in order to evaluate as well as a certain inability to work with the educational team; for these reasons we have called this factor *Non-planned evaluation*. It accounts for a little over 4% of the total variance.

The fifth factor has values in only two questions: it demands an active and compulsory attendance, and requires continuity; for that reason we have called it *Attendance-based evaluation*.

The sixth and last factor is related, mostly, to the use of portfolios or folders, both individual and collaborative, and that is why we have called this factor *Evaluation based on portfolios*; it accounts for 3% of the questionnaire's total variance.

Like we did in the previous scale, we performed a Factorial Confirmatory Analysis (AFC) with the total number of questions of the Evaluation systems scale, and the results of the six factors previously analysed. Both now and then, all indicators of that analysis pointed out to a good set up of the model, and therefore to the existence of those factors we have been talking about (Table VII).

Model	S-B$_{(Chi-square)}$ (gl) (p)	GFI	RMSEA	NFI	NNFI	CFI	AGFI	AIC
Scale factors (six factors)	2687.03 (545)(P = 0.00)	.84	.070	.95	.96	.084	.81	46310.71

Table VII. Evaluation of the six-factor model of the Evaluation systems scale by means of AFC

DISCUSSION AND CONCLUSIONS

The objective of the EMPEF scale to assess how participative methodology and formative evaluation in Physical Education's initial training are perceived by teachers and students has been confirmed by the exploratory factorial analysis.

Results show a seven-factor structure in the case of the methodology and a six-factor structure for evaluation. The absence of other studies to serve as reference that deal with the same aspects as the scale presented here is one that has to be assumed, but it is true that those studies that provide information about methodology and evaluation subjects in general do not delve in as deep as the EMPEF. For instance, the SETLQ (University of Edinburgh, 2001) contains questions referred to goals and congruences (5), elections regarding contents (2), teaching and learning (5), feedback (5), comprehensibility evaluation (2), motivation and eagerness (4), interest shown for the course (2), that do not correspond to the same factors in our study. It is also different from the proposals of Entwistle and collaborators (Entwistle, Hanley, and Hounsell, 1979; Entwistle and Peterson, 2004; Entwistle, Tait and McCune, 2000) with their *Approaches and Study Skills Inventory for Students*, as they tend to use it more for the type of learning achieved (deep, strategic or superficial), although it contains some questions about teaching and evaluation, but they are not related to the concepts of participative methodology and formative evaluation. The work of Gutiérrez-García et al. (2011) does include some items that might relate to the EMPEF, and its population sample is also of initial training in physical education. It contains 17 items for methodology, but it is not clear that they correspond to the same concept of participative methodology exposed here. It does include formative evaluation, and we have to emphasize that the results, for questions that have the same meaning as the EMPEF, show similar results.

To sum up, the EMPEF may be applied by teachers to verify the perception of students regarding participative methodologies and formative evaluation, and to identify the benefits and difficulties of the implementation of innovative conceptions regarding methodology and evaluation. It is ambitious enough to include these aspects of teaching, and has shown the validity needed to be applied in the initial training of Physical Educations degrees. The research lines that may be developed by applying this instrument may help teachers to verify their own teaching and evaluation systems, which would allow them to focus on the indicators they must improve.

REFERENCES

Bain, K. (2005). *Lo que hacen los mejores profesores universitarios*. Valencia: Universidad de Valencia.

Black, P., y Wiliam, D. (2009). Developing the theory of formative assessment. *Assessment, Evaluation and Accountability, 21*(1), 5-31.

Biggs, J. (1987). *Student approaches to learning and studying. Research monograph.* Melbourne: Australian Council for Educational Research. Recuperado de http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=ED308201&ERICExtSearch_SearchType_0=no&accno=ED308201

Biggs, J. (2006). *Calidad del aprendizaje universitario* (2ª ed.). Madrid: Narcea.

Brown, S., y Glasner, A. (Eds.) (2007). *Evaluar en la Universidad. Problemas y nuevos enfoques.* Madrid: Narcea.

Cabrera, A.F., y La Nasa, S. (2002). Sobre los métodos de enseñanza en la Universidad y sus efectos. Diez lecciones aprendidas. En C. Adelman (Ed.), *Nuevas miradas sobre la universidad* (pp. 63-88). Buenos Aires: Universidad Nacional de Tres de Febrero.

Cano-García, F., y Justicia-Justicia, F. (1994). Learning strategies, styles and approaches: an analysis of their interrelationships. *Higher Education, 27*, 239-260.

Cano-González, R. (2009). Tutoría universitaria y aprendizaje por competencias. ¿Cómo lograrlo? *REIFOP, 12*(1), 181-204. Recuperado de http://www.aufop.com

Consejo Europeo (2000). Conclusiones de la presidencia. Recuperado de http://www.europarl.europa.eu/summits/lis1_es.htm

EEES. (2009). Declaración de Leuven/Louvain-la-Neuve, 28-29 April 2009. Recuperado de http://www.eees.es/pdf/Leuven_Louvain-la-Neuve_Communique_April_2009.pdf

Entwistle, N., Hanley, M. y Hounsell, D. (1979). Identifying distinctive approaches to studying. *Higher Education, 8*(4), 365-380.

Entwistle, N. J., y Peterson, E. R. (2004). Conceptions of learning and knowledge in higher education: Relationships with study behaviour and influences of learning environments. *International Journal of Educational Research, 41*, 407-428.

Entwistle, N., Tait, H., y McCune, V. (2000). Patterns of response to approaches to studying inventory across contrasting groups and contexts. *European Journal of Psychology of Education, 15*(1), 33-48.

Feldman, K. A. (1986). The perceived instructional effectiveness of college teachers as related to their personality and attitudinal characteristics: A review and synthesis. *Research in Higher Education, 234*, 129-213.

Fernández Pérez, M. (1989). *Así enseña nuestra Universidad: Hacia la construcción crítica de una didáctica universitaria.* Madrid: Editorial Complutense.

Fuentes-Medina, M. E., y Herrero, J. R. (1999). Evaluación docente: Hacia una fundamentación de la autoevaluación. *Revista Electrónica Interuniversitaria de Formación del Profesorado, 2*(1). Recuperado de http://www.uva.es/aufor/publica/revelfor/99-v2n1.htm

Gimeno, J. (2012). ¿Por qué habría de renovarse la enseñanza en la Universidad? En J.B. Martínez (Coord.), *Innovación en la universidad. Prácticas, políticas y retóricas* (pp. 27-51). Barcelona: Graó.

Gutiérrez-García, C., Pérez-Pueyo, A., Pérez-Gutiérrez, M., y Palacios-Picos, A. (2011). Percepciones de profesores y alumnos sobre la enseñanza,

evaluación y desarrollo de competencias en estudios universitarios de formación de profesorado. *Cultura y Educación, 23*(4), 499-514.

Huber, G.L. (2008). Aprendizaje activo y metodologías educativas. *Revista de Educación, número extraordinario*, 59-81.

Irons, A. (2008). *Enhancing Learning through Formative Assessment and Feedback*. New York: Routledge.

Knight, P.T. (2005). *El profesorado de educación superior*. Madrid: Narcea.

Latiesa, M. (2000). Validez y fiabilidad de las observaciones sociológicas. En M. García Ferrando, J. Ibáñez, y F. Alvira (Comps.), *El análisis de la realidad social. Métodos y técnicas de investigación* (pp. 409-443). Madrid: Alianza.

Leathwood, C., y Phillips, D. (2000). Developing curriculum evaluation research in higher education: Process, politics and practicalities. *Higher Education, 40*, 313-330.

Losada, J. L., y López-Feal, R. (2003). *Métodos de investigación en ciencias humanas y sociales*. Madrid: Thomson.

McMillan, J. H. (ed.) (2007). *Formative Classroom Assessment. Theory into Practice*. New York: Teachers College Press.

Marín-García, J. A., Miralles-Insa, C., García-Sabater, J. J., y Vidal-Carreras, P. I. (2008) Enseñando Administración de Empresas con docencia basada en el trabajo en equipo de los estudiantes: Ventajas, inconvenientes y propuestas de actuación. *Intangible Capital, 4*(2), 143-165.

Marsh, H. W. (1982). SEEQ: A reliable, valid, and useful instrument for collecting students' evaluations of university teaching. *British Journal of Educational Psychology, 52*, 77-95.

Marsh, H. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research, 11*, 253-388 (whole issue).

Marsh, H. W. (2007). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases, and usefulness. En R. P. Raymond, y J. C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 319-383). Dordrecht ,The Netherlands: Springer

Marsh, H. W., y Roche, L. A. (1997). Making students' evaluations of teaching effectiveness effective. The critical issues of validity, bias, and utility. *American Psychologist, 52*(11), 1187-1197.

Meyer, C., y Jones, T. B. (1993). *Promoting active learning: Strategies for the college classroom*. San Francisco: Jossey-Bass.

Millis, B. (ed.) (2010). *Cooperative Learning in Higher Education: Across the Disciplines, Across the Academy*. Sterling: Stylus Publishing.

Murray, H. G., Rushton, J. P., y Paunnem, S. V. (1990). Teacher personality traits and student instructional ratings in six types of university courses. *Journal of Educational Psychology, 82*(2), 250-261.

Palacios, A. y López-Pastor, V. (2013). Haz lo que yo digo pero no lo que yo hago: sistemas de evaluación del alumnado en la formación inicial del profesorado. *Revista de Educación*, *361*, 279-305.

Pérez, Á., Julián, J. A., y López-Pastor, V. M. (2009). Evaluación formativa y compartida en el Espacio Europeo de Educación Superior. En V.M. López-Pastor, (coord.), *Evaluación formativa y compartida en Educación Superior* (pp. 19-43). Madrid: Narcea.

Prieto, L. (2007). *Autoeficacia del profesor universitario: eficacia percibida y práctica docente*. Madrid: Narcea.

Race, P. (2007). *The Lecturer's Toolkit: A practical guide to assessment, learning and teaching*. New York: Routledge.

Ramos, F., y Ortiz, M. (2010). Una experiencia de auto-evaluación de profesorado: teoría y práctica. *Exedra, 4*, 81-107.

Ramsden, P., y Entwistle N.J. (1981). Effects of academic departments on students' approaches to studying. *British Journal of Educational Psychology, 51*(3), 368-383.

Stenhouse, L. (1987). *La investigación como base de la enseñanza*. Madrid: Morata.

Tejedor, F. (Dir.) (1998). *Las estrategias utilizadas por los profesores universitarios para la evaluación del aprendizaje de los alumnos*. Madrid: CIDE (informe).

Tejedor, F. (2003). Un modelo de evaluación del profesorado universitario. *Revista de Investigación Educativa, 21*(1), 157-182.

Tippin, G.K., Lafreniere, K.D., y Page, S. (2012). Student perception of academic grading: Personality, academic orientation, and effort. *Active Learning in Higher Education, 13*(1), 51-61.

Trillo, F. (2005). *La evaluación de los estudiantes universitarios. El caso de la Universidad de Santiago de Compostela*. Santiago de Compostela: Universidad de Santiago de Compostela.

Trillo, F., y Porto, M. (1999). La percepción de los estudiantes sobre su evaluación en la universidad. Un estudio en la Facultad de Ciencias de la Educación. *Innovación Educativa, 9*, 55-75.

University of Edinburgh (2001). ProjectETL. Shortened Experiences of Teaching and Learning Questionnaire (SETLQ). Disponible en www.etl.tla.ed.ac.uk/docs/SETLQscoring.pdf

Vázquez, J. A. (2008). La organización de las enseñanzas de grado y postgrado. *Revista de Educación, número extraordinario*, 23-39.

Villa, A. (2008). La excelencia docente. *Revista de Educación, número extraordinario*, 177-212.

Referencias totales / Total references: 48 (100%)
Referencias propias de la revista / Journal's own references: 0 (0%)

QUESTIONNAIRE ON PARTICIPATIVE METHODOLOGY AND FORMATIVE ASSESSMENT

With the aim of studying the methodology and evaluation systems used for the subject, we ask you to answer sincerely to the following questions. The questionnaire is ANONYMOUS, and we guarantee its CONFIDENTIALITY. Thanks for your help

Finished courses	Secondary school	Professional training	Certificate	Degree	Other

Gender	Male	Female	Age		Number of times you have registered for this subject	

Approximate percentage of the subject's classes you have attended	I have not attended any classes	Less than 25%	Between 25 and 50%	Between 50 and 75%	Between 75 and 90%	Over 90%

1.- Was the subject matter's program negotiated at the beginning of the course?	No/Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply

2.- Has the methodology used in this subject helped you acquire professional competences?	No/Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply

3.- Does the assessment outlined favour the acquisition of professional competences?	No/Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply

4.- To what extent have the following methodologies and strategies been used throughout the subject?

	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
01. Tutelary projects	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
02. Seminar	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
03. Case studies	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
04. Learning based on problems / problem solving	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
05. Collaborative learning	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
06. Guest speech	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
07. Dialogic meetings	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
08. Debates	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
09. Checking bibliography	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
10. Observation of teaching practices	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
11. Round tables	Not	A	To some	Quite a	A lot	No opinion/No

	at all	little	extent	bit		reply

5.- To what extent did you use in the subject the following instruments?

01. Field notebooks	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
02. Journal or session sheets	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
03. Books or articles reviews	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
04. Reports	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
05. Virtual platforms	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
06. Notebook or similar to write down notes	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply

6.- Point out how many times were the following statements true:

01. A positive classroom climate was favoured	Never	Few times	Sometimes	Quite a few times	Many times	No opinion/No reply
02. Questions were allowed in classes	Never	Few times	Sometimes	Quite a few times	Many times	No opinion/No reply
03. Tutorials were used to solve doubts about the subject's progress	Never	Few times	Sometimes	Quite a few times	Many times	No opinion/No reply
04. Audiovisual media have been used in classes (transparencies, presentations, videos, etc.)	Never	Few times	Sometimes	Quite a few times	Many times	No opinion/No reply
05. The subject was accompanied by practices that helped to understand better the future professional work	Never	Few times	Sometimes	Quite a few times	Many times	No opinion/No reply
06. A variety of methodological strategies was used	Never	Few times	Sometimes	Quite a few times	Many times	No opinion/No reply
07. Oral expositions were used in classes to give explanations, and notes were taken	Never	Few times	Sometimes	Quite a few times	Many times	No opinion/No reply

7.- Regarding the overall assessment system, indicate how many times were the following statements true during the subject's duration:

01. Continuous assessment processes were performed (activities or documents assessments during the subject's duration)	Never	Few times	Sometimes	Quite a few times	Many times	No opinion/No reply
02. Formative assessment processes were used (the teacher corrected activities or documents, and provided feedback as to how to improve and correct mistakes)	Never	Few times	Sometimes	Quite a few times	Many times	No opinion/No reply
03. The observations made regarding essays or required activities might result in their repetition in order to improve them	Never	Few times	Sometimes	Quite a few times	Many times	No opinion/No reply
04. A portfolio or individual folder was used to hand in materials	Never	Few times	Sometimes	Quite a few times	Many times	No opinion/No reply
05. A group portfolio or collaborative folder was used to hand in materials	Never	Few times	Sometimes	Quite a few times	Many times	No opinion/No reply

	Never	Few times	Sometimes	Quite a few times	Many times	No opinion/No reply
06. Quality criteria for different activities, essays or exams were previously explained	Never	Few times	Sometimes	Quite a few times	Many times	No opinion/No reply
07. There were similarities with the assessment systems of other subject matters in this same centre	Never	Few times	Sometimes	Quite a few times	Many times	No opinion/No reply

8.- Regarding the type of exam, indicate whether you used:

	Yes	No	No opinion/No reply
01. Oral exam	Yes	No	No opinion/No reply
02. Written exam with long questions developing a certain theme or similar	Yes	No	No opinion/No reply
03. Exam with short questions	Yes	No	No opinion/No reply
04. Multiple choice test	Yes	No	No opinion/No reply
05. Practical exam (solving problems, proposing activities, etc.)	Yes	No	No opinion/No reply
06. There was no exam	Yes	No	No opinion/No reply

9.- Regarding the quality of the subject, indicate whether:

	Yes	No	No opinion/No reply
01. The grade depends only on the final exam	Yes	No	No opinion/No reply
02. The final grade depends on the final exam and an individual essay	Yes	No	No opinion/No reply
03. The grade depends on the exam and other instruments (notebooks, reviews, group work, etc.)	Yes	No	No opinion/No reply
04. The final grade is the sum of the grades of the partial exams taken	Yes	No	No opinion/No reply
05. The grade is obtained without exams	Yes	No	No opinion/No reply
06. The grading system and criteria are explained at the beginning of the subject and are well known	Yes	No	No opinion/No reply
07. The grade assigned by the teacher is justified	Yes	No	No opinion/No reply
08. In spite of preparing different theoretical-practical essays, the exam had a definitive weighting on the final grade (if the exam was not passed, neither was the subject)	Yes	No	No opinion/No reply
09. There have been theoretical-practical essays (not exams) that have greatly influenced the final grade (subject was not passed if they were not passed)	Yes	No	No opinion/No reply

10.- Regarding the participation of students in the assessment and grading process of the subject, indicate whether:

	Yes	No	No opinion/No reply
01. Students made an assessment of their own work (self-evaluation)	Yes	No	No opinion/No reply
02. Students assess their classmates' work (co-evaluation)	Yes	No	No opinion/No reply
03. The final grade is agreed by both teacher and student (grade by consensus)	Yes	No	No opinion/No

					reply
04. Students themselves assigned a grade in a justified manner (self-grade)	Yes	No	No opinion/No reply		

11.- To what extent do you feel the subject develops the following professional competences?

01. Organizing and favouring learning situations	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
02. Managing the progression of learning	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
03. Preparing and implementing strategies to take into account diversity	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
04. Involving students in the learning process and academic life	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
05. Teamwork with other teachers	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
06. Taking part in the teaching centre's management	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
07. Informing and involving families	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
08. Using information/communication technologies	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
09. Facing the duties and ethical dilemmas of the teaching profession	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
10. Establishing formative evaluation processes	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply

12.- Indicate whether you agree or not with the following statements regarding the subject's assessment system:

01. It offers alternatives for all students	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
02. There is a previous contract, agreed and negotiated, regarding the assessment system	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
03. It is centred on the process, the importance of daily work	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
04. The student performs an active learning	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
05. Teamwork is conceived in a collaborative manner	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
06. The student is more motivated, and the learning process is more motivational	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
07. Grades are more fair	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
08. Improves academic tutelage (follow-up and help for students)	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
09. Allows functional learnings	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
10. Generates significant learnings	Not at all	A little	To some	Quite a bit	A lot	No opinion/No

	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
			extent			reply
11. Much more is learnt	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
12. Improves the quality of requested essays	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
13. There is a correlation between theory and practice	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
14. Evaluates all possible aspects	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
15. There is feedback and the possibility to correct mistakes in essays and activities	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
16. A more personalized follow-up is provided	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
17. Requires more responsibility	Nothing	A little	To some extent	Quite a bit	A lot	No opinion/No reply

13.- Indicate whether you agree or not with the following statements regarding the subject's assessment system:

	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
01. Demands compulsory and active attendance	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
02. Its work dynamic is little known, lack of custom	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
03. Demands continuity	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
04. It must be previously understood	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
05. Demands a greater effort	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
06. It is difficult to work in teams	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
07. A lot of work may be accumulated towards the end	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
08. The relationship work/credit is disproportionate	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
09. The process is complex and sometimes unclear	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
10. Generates uncertainty and insecurity, doubts about what is to be done	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
11. It is unfair compared to other assessment processes	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
12. Corrections were not clear enough	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
13. The assessment of the work done is subjective	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply
14. It demands to participate in my own assessment (self-evaluation)	Not at all	A little	To some extent	Quite a bit	A lot	No opinion/No reply

14.- Indicate your overall satisfaction regarding the subject

Not at all satisfied	A little satisfied	Reasonably satisfied	Quite satisfied	Very satisfied	No opinion/No reply

15.- Indicate your overall satisfaction regarding the subject's assessment

Not at all satisfied	A little satisfied	Reasonably satisfied	Quite satisfied	Very satisfied	No opinion/No reply

16.- What is the level of difficulty of the subject?

Not at all difficult	Not very difficult	Reasonably difficult	Quite difficult	Very difficult	No opinion/No reply

Would you like to make any comments regarding the subject or this questionnaire? (You may use the space you deem convenient either here or in the other side of the page)

<div align="center">

-THANKS FOR YOUR HELP-

</div>