The Qualitative Report

Volume 29 | Number 6 Article 12

6-12-2024

Creating Assessment Rubrics for Final Teacher Education Degree Projects: A Qualitative Case Study

Carla Fernández Garcimartín
University of Valladolid, carlafdzg96@gmail.com

Víctor Manuel López Pastor
University of Valladolid, victor.lopez.pastor@uva.es

Teresa Fuentes Nieto
University of Valladolid, fuentes.teresa@gmail.com

David Hortiguela Alcalá
University of Burgos, dhortiguela@ubu.es

Recommended APA Citation

Fernández Garcimartín, C., López Pastor, V. M., Fuentes Nieto, T., & Hortiguela Alcalá, D. (2024). Creating Assessment Rubrics for Final Teacher Education Degree Projects: A Qualitative Case Study. *The Qualitative Report*, *29*(6), 1689-1711. https://doi.org/10.46743/2160-3715/2024.6426

Creating Assessment Rubrics for Final Teacher Education Degree Projects: A Qualitative Case Study

Abstract

This paper analyzes the process of creating the assessment instruments for the Final Year Project (FYP) by education faculty based on the opinions of the lecturers. Qualitative methodology was used, with a case study design and the in-depth interview as a data collection technique. Four lecturers participated as relevant informants. The results show that: (a) the faculty decided to create the instruments due to the lack of clear assessment criteria among the lecturers in the first years (2009-2015); (b) work was carried out on three assessment instruments over two years (2015-2017): a tutor's report (without grade), a document rubric (80% of the mark), and an oral presentation rubric (20% of the mark). It is concluded that the assessment instruments were created through the necessity of the faculty to achieve a more objective and effective FYP assessment. The creation of the instruments followed some steps to encompass the degree competencies and the objectives, types, and criteria of the FYP. The paper presents the process of creating assessment instruments and the advantages of their use, specifically the treatment of rubrics as the main assessment instrument. It seems necessary to carry out further research on assessment instruments for dissertations and theses, given their proven relevance.

Keywords

case study, assessment, rubric, grade, project, pre-service teacher education

Acknowledgements

This research was supported by Grant RTI2018-093292-B-I00 funded by MCIN/AEI/ 10.13039/ 501100011033 and by "ERDF A way of making Europe."

The Qualitative Report 2024 Volume 29, Number 6, 1689-1711
https://doi.org/10.46743/2160-3715/2024.6426

Creating Assessment Rubrics for Final Teacher Education Degree Projects: A Qualitative Case Study

Carla Fernández-Garcimartín[1], Víctor Manuel López-Pastor[1], Teresa Fuentes Nieto[1], and David Hortigüela-Alcalá[2]
[1]University of Valladolid, Spain
[2]University of Burgos, Spain

This paper analyzes the process of creating the assessment instruments for the Final Year Project (FYP) by education faculty based on the opinions of the lecturers. Qualitative methodology was used, with a case study design and the in-depth interview as a data collection technique. Four lecturers participated as relevant informants. The results show that: (a) the faculty decided to create the instruments due to the lack of clear assessment criteria among the lecturers in the first years (2009-2015); (b) work was carried out on three assessment instruments over two years (2015-2017): a tutor's report (without grade), a document rubric (80% of the mark), and an oral presentation rubric (20% of the mark). It is concluded that the assessment instruments were created through the necessity of the faculty to achieve a more objective and effective FYP assessment. The creation of the instruments followed some steps to encompass the degree competencies and the objectives, types, and criteria of the FYP. The paper presents the process of creating assessment instruments and the advantages of their use, specifically the treatment of rubrics as the main assessment instrument. It seems necessary to carry out further research on assessment instruments for dissertations and theses, given their proven relevance.

Keywords: case study, assessment, rubric, grade, project, pre-service teacher education

Implementation of Final Year Projects in University Studies

In Spain, the Bologna Declaration (1998) gave rise to the "Bologna Plan" (1999), which proposes that university students should develop general, specific, critical, and creative skills within a professional framework (Royal Decree, 1393/2007). The aforementioned Royal Decree states that all undergraduate studies will be considered completed once a "Final Year Project" (FYP) has been prepared and defended. Mateo et al. (2012) define the FYP as a project developed and carried out by a student upon finishing his/her academic degree that should demonstrate the competencies acquired throughout their studies, including: autonomy, initiative, and problem-solving strategies in their field. Jawitz et al. (2002) as well as Sharef et al. (2014), define the FYP as the final academic step to finishing a student's course of study, most likely the most important one in the whole degree, the quality of which can be taken as an indicator of the knowledge acquired. Hashim and Hashim (2010) add that the FYP contributes to: (1) problem solving in research through techniques, tools, skills, etc.; (2) the design, analysis and assessment of research projects; and (3) the effective presentation of study results and development of a formal, technical document. Medina et al. (2020) suggest that the FYP furthers the development of skills and competencies learned during undergraduate studies.

Royal Decree (1393/2007) gives the universities autonomy to elaborate the criteria for developing and assessing the FYP. García and Valle (2014) add that it is up to university lecturers to create a precise FYP development guide, which adapts the European regulations on official university education to the characteristics and needs of their university. Each university manages its regulations in accordance with the aforementioned Royal Decree, providing regulations adapted to its context and mandatory for university professors.

The preparation of an FYP involves not only the work of the student but also that of part of the degree teaching staff. Thus, two parties and the students are present in the preparation and assessment of the FYP: the tutor and the assessment committee.

The tutor is essential in the correct preparation of the FYP. Tutors are chosen by the students according to the FYP topic they are leading or according to the student's average mark in the degree. Vera and Briones (2014) find that students highlight tutoring as a key element for the correct development and learning of the work. Sancho-Esper (2018) specifies the tasks of the tutor during the elaboration of the FYP: to monitor the evolution of the work, guide the student, and ensure that the objectives set are met.

The assessment panel also plays an important role in the feedback, assessment and grading. Depending on the university, there are two or three professors from the same faculty, outside the FYP. Vicario-Molina et al. (2020) state that it is made up of lecturers in charge of assessing the development and grading of both the final document and the oral presentation of the FYP.

Assessment of Final Degree Projects

All the agents involved in the assessment (Degree student, tutor and assessment committee) should evaluate the process of the work from the beginning to the end, since it seems complicated to objectively rate a project just by reading the final product. Vicario-Molina et al. (2020) consider the assessment of the FYP as a key item to evaluate the student's training in knowledge and skills related to their training.

Quintana and Gil (2015) state that in order to develop the FYP it is necessary to have a continuous assessment system that works as "graduation assurance" (p. 2), showing the students what they should achieve and reporting their learning and evolution, so that they are able to identify their achievements and aspects to improve. Therefore, the formative assessment system would be the most coherent assessment model to use during the development of the FYP, understanding it as an assessment system in which the lecturer gives feedback to the students in their learning process and allows them to orient themselves within the task and recognize strengths and weaknesses (Marín et al., 2015).

Assessment instruments are needed both for evaluating and grading FYPs. López-Pastor and Pérez-Pueyo (2017) define assessment instruments as a set of tools that assess the degree of achievement of competencies and provide information throughout the teaching-learning process, and Quintana and Gil (2015) present the characteristics they should have. The FYP assessment tools should include everything that is expected from the student, including the completion of the work and clear assessment criteria and grading criteria.

From a practical point of view, a complete and thorough assessment of all the general competencies of a degree with a FYP is complex and far-removed from realistic application. Although few studies have been done on the assessment of FYPs, within them, measurement of competencies has, in fact, been highlighted (Hashim & Hashim, 2018; Medina et al. 2020). Since the elaboration of a FYP entails demonstrating the knowledge and skills required for an undergraduate degree, one could assume that the FYP assessment criteria must reflect achievement of those competencies. Taking that into account, De Sande et al. (2011) assert that assessing all competencies in a FYP is extremely complex, especially since some projects do

not require them all. Precisely in that same line, Medina et al. (2020) clarify that although a FYP contributes to the assessment of the competencies of a degree, that should not mean all competencies must be evaluated in every project. It should be seen as contributing to the outcome.

Assessment Instruments for FYP in Faculties of Education

The assessment of the FYP should be in line with the assessment in pre-service teacher education. Ruiz (2015) indicates that within Pre-Service Teacher Education (PTE) the assessment of the FYP should fulfill the following objectives: (a) certify the learning required in the degree; (b) favor the achievement of learning, and (c) help the student's personal growth. In recent years, using the term "rubric" to refer to "descriptive scales" has become very common. López-Pastor and Pérez-Pueyo (2017) describe rubrics as scales that establish different grades or levels, each with precise descriptions of the characteristics a student's production may have or the possible behaviors or responses to a learning activity.

According to Brookhart (2018), there are two types of rubrics: holistic and analytic. Holistic rubrics are made up of general, not detailed, assessment criteria, while analytic rubrics are thorough, detailed, and tend to be more complex. This author explains how no evidence suggests what type of rubric is better. However, analytical rubrics give a more detailed analysis of performance.

The research reviewed in this study affirms the effectiveness of rubrics as a main instrument of assessment, although this could be called into question when considering whether or not they are correctly employed in a formative and transparent way. In this regard, Panadero and Jonsson (2013) and Sharef et al. (2014) argue that rubrics are useful for clarifying the project's objectives and competencies and the formative scope of learning. Furthermore, these authors add that rubrics may have the potential to assess formatively and teach through a student-centered approach to assessment. Black and William (2009) argue that projects can be considered formative in class when both students and lecturers obtain, analyze, and employ instruments that boost the learning progress and student performance.

Marín et al. (2012) state that rubrics serve to: (1) clarify the student's formative process; (2) show the strengths and weaknesses acquired throughout the teaching-learning process; (3) promote student responsibility; (4) foster critical thinking and self-knowledge of learning; (5) systematize and compile information and evidence of student work; and (6) increase student motivation and the formative nature of the instrument. In the same logic, Ruiz (2015) defends the idea that the rubric makes formative sense if the student knows from the first moment what is going to be assessed with the assignment. Hattie and Timperley (2007) and Brooks et al. (2019) explain that the feedback given to the student will have a greater or lesser impact on learning depending on the form it offers. In this regard, Hattie and Timperley (2007) state that feedforward is used to improve on the task performed, but with a view to the future, and has the following characteristics: (a) it helps the learner to self-regulate more effectively in the face of learning; (b) it helps to obtain more strategies to deepen the understanding of what is being learned.

Table 1 shows two different processes of elaboration of descriptive scales.

Table 1
Steps for developing descriptive scales/assessment rubrics

Mertler (2001)	López-Pastor & Pérez-Pueyo (2017)
1-Consider the context and the general conception of the task: set objectives, competencies and evidences of the activity.	1-Choose the learning objectives and contents to be achieved with the task.
2-Describe the criteria to correctly elaborate the task and group them into dimensions.	2-Identify the specific learning activity and select the aspects that will need to be developed.
3-Design the rating scale: decide on the levels of achievement, write the descriptors for each level and name them.	3-Define clearly and synthetically the levels of achievement of the development of the task.
4-Select work samples from each performance level.	4-Determine the type of assessment for each level of achievement: qualitative or quantitative.
	5-Evaluate the design of the instrument through another instrument.
	6-Apply the instrument as a test and evaluate its suitability.
	7-Put the instrument into practice in a real situation and reflect with the students on its usefulness, suitability and clarity.

Source: Own elaboration based on Mertler (2001) and López-Pastor and Pérez-Pueyo (2017).

No studies have been found on the development and use of rubrics for the assessment of final year projects in Spanish faculties of education. We can only take as a specific reference Huaman (2015), who proposes a sequence of eight steps to elaborate a rubric in a faculty of education in Peru: (1) presentation of subject data; (2) determination of the competencies that students will work on; (3) exposition of criteria and indicators related to each competency; (4) elaboration of the levels of mastery of the competencies; (5) elaboration of the descriptors of each level; (6) determination of the quantitative value of each indicator; (7) adequacy of the rubric as self-assessment, co-assessment and hetero-assessment; and (8) analysis of the rubric by the students before performing the task.

There is a considerable lack of studies on the subject of rubrics as assessment tools for dissertations and theses in Spanish faculties of education. For this reason, the objectives of this study are the following:

> 1-To review the procedure and criteria used in the assessment of the FYP in a Spanish faculty of education between the years 2009-2015, when there were no specific instruments for this purpose.

2-To analyze the reasons and the process of creation of the FYP assessment instruments in 2015 in that faculty.

Context

The study is framed within a public university located in the center of a city near the capital. It is a small environment with approximately 500 students in the Teacher Education degree. Students taking the FYP take it as the last subject of their degree and are usually around 22 years old. The FYP is one of the most important subjects of the degree, since the student must demonstrate the competences acquired during the years of academic study in a theoretical or theoretical-practical project.

This research is carried out by professors and doctors from a Spanish university, within the degree of education. They work in the line of research on formative assessment, focusing on final degree projects.

Method

Design

In this case, we are dealing with a qualitative research study. According to Krause (1995), these are procedures that make possible a construction of knowledge that occurs on the basis of concepts that reduce the complexity of the study by establishing relationships between them and generating internal coherence. Specifically, methodologically, an interpretative qualitative paradigm is approached that conceives the research process inductively, where categories or themes develop from informants, rather than anticipating them from the outset (Ceballos-Herrera, 2009).

A case study was carried out. Stake (1995, p. 11, 2002) defines it as: "the study of the particularity and complexity of a singular case, in order to understand its activity in important circumstances." The case study is the most appropriate research design for the research objectives. In relation to the case study in question, an "intrinsic" type of case study is used. According to Álvarez-Álvarez and San Fabián-Maroto (2015), cases are analyzed with their own specificities, which have a value in themselves and aim to achieve a better understanding of the specific case to be studied with an interest in itself. This is so because they focus on analyzing what happened in a very specific reality, a faculty of education in a small Spanish city, and a very specific aspect: how teachers organized themselves to assess FYP in that faculty when the change of curricula introduced this type of subject for the first time as a compulsory final project to complete university studies.

Participants

Four lecturers from a Spanish faculty of education participated. They were selected as key informants for different reasons: (a) all participated in the process of creating the FYP assessment instruments; (b) three lecturers because they coordinated the whole process of elaborating the new assessment instruments and defend the idea that having "descriptive scales" (rubrics) allows for a more reliable and consensual assessment; (c) one lecturer because he shows a more critical stance with the general use of the instruments in all types of FYP. In this way, different perspectives can be observed which help to understand the process from different points of view and pedagogical reasons. They were informed of the aim of the study and asked for consent to participate in the study, with guarantees of anonymity, confidentiality

and review of draft transcripts. Table 2 presents the characteristics of each subject and the codes used in the data analysis.

Table 2
Codes and description of the subjects participating in the study and named in the in-depth interviews.

CODE	SUBJECT	DESCRIPTION
C1-C2-C3	Lecturer 1,2 & 3	Coordinators of the project for the creation of the FYP assessment instrument (2015-2017).
P1	Lecturer 4	Participant in the process of creation of the instruments (2015-2017). Shows a critical stance with the use of rubrics, as opposed to the majority of the faculty.

Source: Own elaboration.

Data collection techniques

In-depth interviews were conducted with the subjects indicated in Table 2. They were carried out in a semi-structured manner, as set out by Krause (1995), starting from the basis of the questions elaborated in the initial script and being flexible in it; thus, it was possible to obtain richer and more personal information from each lecturer. The initial script had generic questions, which we then fleshed out in a delimited interview format. Maxwell (2019) argues that this technique is employed to understand the personal perspectives of the research subjects.

A single interview was conducted with each participant, lasting approximately one hour. The following are examples of questions drawn from the interviews:

- How did you assess and grade FYPs before these rubrics you created existed?
- What criteria did you use to evaluate and then grade FYPs, being an assessment board?
- What was the motivation for creating the rubrics?
- How did you organize yourselves as Teacher Education teachers?
- Were students involved in the process of creating them?

After the interview, the data were transcribed and sent to each participant (interviewee) for modification and validation.

In our study, the starting data are the transcripts of the interviews conducted, after having been reviewed, modified and validated by the four participants. This is explained in more detail in the following section.

Procedure

An analysis of the categories related to the research objectives was made, which aided in the investigation and collection of specific results (Vaismoradi et al. 2016). Krause (1995) and Creswell (2013) define a category as a unit of information composed of different occurrences and argue that for the analysis of a qualitative study by categories, we must first code the data and then categorize it. In line with Zhang and Wildemuth (2009), the analysis was based on the theoretical aspects of the research. The information must be fragmented to code the data, thus dividing the data obtained. The categories are directly related to the two research objectives. The subcategories, on the other hand, emerge from a first review of the

data, when we check which are the topics most frequently dealt with by the teachers interviewed. With this first system of categories and subcategories, all the units of meaning obtained from the transcriptions of the interviews are coded. A "unit of meaning" is the sentence or part of a sentence that talks about a topic that clearly corresponds to one of the established subcategories; this is the basic unit of analysis to be coded. Subsequently, the subcategories are checked to ensure that they are mutually exclusive and that they correctly fulfil the criteria of saturation and appropriateness. Each literal quotation from the interviewees will be followed by the code of the corresponding interviewee (see Table 2).

Table 3 shows the system of categories obtained from the data analysis.

Table 3
Selected categories and subcategories

CATEGORY	SUBCATEGORIES
1-Assessment of the FYP before the creation of the assessment instruments (2009-2015)	1.1.-Procedures for the assessment of the FYP before the creation of the instruments (2009-2015). 1.2.-Criteria for the assessment of the FYP before the creation of the instruments (2009-2015).
2-Process for the creation of the FYP assessment instruments (2015-2017).	2.1. Motivation to create the instruments and types. 2.2. Process of creation of the instruments. 2.3. Final process of experimentation.

Below are several examples from the interview transcripts that correspond directly to the codes presented in Table 3:

> There were no objective criteria. The contents, the development of the objectives, what they were, if they were well developed, if they had been achieved and so on, the conclusions. (1.1)

> We had a small document, simply to guide us a little. But, in general, there was nothing. (1.2)

> From 2007 there was certainly nothing. Well, there was a reference guide describing, according to the 2007 Decree itself, what the purpose of a FYP was and some basic indications of quality. From there, a Teaching Guide was generated as for any other subject, which was posted on the Virtual Campus; a teaching project, in this case, always saying “aspects such as” will be assessed. (1.2)

> We needed an assessment instrument, something homogeneous, that would be useful to all of us. (2.1)

> Well, we can say that there was a kind of small coordinating committee, where we would be talking about four or five teachers carrying a little bit of the weight and leadership of the meetings, but the truth is that, in general, there was a lot of participation because every time we held a meeting, we could talk about 22 or 23 teachers in a seminar discussing rubrics. (2. 2)

> It was even suggested that there were different types of rubrics, but the truth is that the heterogeneity of the participating teaching staff encouraged us to have an analytical rubric (laughs), because if the precedents had been those of a fairly similar assessment, even between committees and even very similar criteria, then there would have been no need for an analytical rubric. But in this case, as there were discrepancies, it was the best solution. (2.3)

Throughout the article, the “descriptive scales” will be referred to as "rubrics" because this is the term most commonly used in scientific articles and the interviewees' interventions.

To ensure the scientific rigor of the study, we have based ourselves on the criteria elaborated by Guba (1989) for naturalistic studies, with qualitative methodology:

- Credibility: transcribing and analyzing the qualitative data as they were collected, with the consent of each subject. Transcriptions were reviewed by the interviewees, with the possibility of modifying aspects that did not reflect their thinking. This increases the validity and reliability of the process and results; conducting a second round of interviews to validate specific data that were not clear in the first round.
- Transferability: The procedure followed and the analysis of possible aspects that could be transferable to similar realities, of which there are many in our country, are described. It is considered that "transferability" is a basic characteristic in qualitative research, especially when dealing with single cases and very small samples, as is the case in this study. Therefore, the results of the research are not directly applicable to other contexts, but they can be easily transferable to similar contexts.
- Dependence: triangulation of data collection techniques and subjects, since the analysis of categories and subcategories is verified with each participant. As a result, by being objective as researchers, it has been possible to guarantee the stability, the coherence of the analysis and the replicability of the study in similar contexts. The analysis carried out was completely objective, avoiding intersubjective researcher bias.
- Confirmability: triangulation of techniques and subjects, as well as rigorous data analysis. The intersubjectivity of the participants is carried out on and the researchers' opinion bias is avoided; so, the researchers have analyzed the results obtained in an objective manner without interjecting their opinion on the results obtained.

The results follow the same structure as Table 3 of categories. This section is organized in two parts, giving rise to a large section containing the "categories" and sub-sections within them corresponding to the "sub-categories."

Results

The results follow the same structure as Table 3 of categories.

Assessment of the FYP before the creation of the assessment instruments (2009-2015)

This first section of results is organized into two sub-sections, which correspond to the system of categories and sub-categories indicated above: (a) assessment procedures of the

FYPs before the creation of the instruments; (b) Assessment criteria of the FYPs before the creation of the instruments.

Assessment procedures of the FYPs before the creation of the instruments

In this section we analyze what the FYP assessment process was like before the creation of the specific rubrics that were designed and used from 2015 onwards. Taking as a reference the general guide for the development and assessment of the FYP of Royal Decree (1393/2007), this faculty of education drew up a specific document in which were developed: (a) general aspects of the assessment of the FYP; (b) the specific assessment criteria of the work; (c)the competences of the tutor and the assessment committee. The data analysis shows that the guide is not really considered and that when it comes down to it, there are no common criteria, but that each teacher uses the assessment criteria he or she considers appropriate. Lecturers elaborated this document in order to clarify and contextualize this guide in their faculty. In theory, this document was very important for making decisions about the assessment of the FYP, but in reality, it was little used, as one of the coordinators acknowledges:

> From 2007 there certainly wasn't. Well, there was a reference guide describing, based on the 2007 Decree itself, what the purpose of a FYP was and some basic indications of quality. From there, a Teaching Guide was generated as for any other subject, which was posted on the Virtual Campus; (...) always saying "aspects such as" will be assessed. (C2)

The FYP evaluation process has three aspects to take into account: (a) the tutor does not assess, he/she only makes a report to the panel on the student's work during the elaboration of the FYP, but without grading; (b) the tribunal, initially composed of 3 people, is in charge of assessing and marking the FYP; or (c) the committees reach the final grade in different ways; sometimes through a debate on the quality of the work with a final consensus, but sometimes through calculating the average between the three grades. Given that the time dedicated to the defense and evaluation of each FYP is limited, in many cases the most commonly used solution is to calculate the average, when they see that they do not reach a consensus on the evaluation criteria. The following quotes refer to these aspects and how to arrive at the final grade:

> The tutor did not assess anything; it was evaluated by the tribunal. There were three people, each one gave his grade, and we made an average. There was no consensus (...). Each one had read it at home, then we came and, (...) once we discussed with the student, we tried to reach a consensus. It was not so much about saying: "Well, come on, let's do the average." (C1)

> Each member of the assessment committee, depending on those large blocks of criteria, had their notes or annotations and, in addition, a grading proposal. The only thing that was done was to discuss (...) to reach an agreement on what grade to put based on what strong and weak aspects of the work. (C2)

Therefore, it seems to be understood that there were different assessment procedures, depending on the people who formed the assessment committee and the limited time available for each FYP. In some cases, the final rating was discussed on the basis of criteria (even if they were not the same) and in some cases a technical solution of coexistence or temporary efficiency was chosen: the average of the different ratings. The time limits had a considerable impact on this type of process. In the course of a morning, each panel had to assess between 8

and 10 FYPs, and only 20-25 minutes were available for the student to defend the FYP (10 minutes), discuss with the committee, and then, behind closed doors, the committee decided the final grade and passed it on to the student and the tutor. In many cases, this time pressure led to the use of the solution of an arithmetic mean solution, if the teachers of the committee did not reach a quick agreement.

Criteria for the assessment of the FYP before the creation of the instruments (2009-2015)

The criteria for assessing the FYP from 2009 to 2015 appeared in the specific faculty document and marked hints on how the assessment of the FYP should be. One coordinator considers that these criteria were not objective and another that they were not applied or were poorly applied:

> There were no objective criteria: the contents, the development of the objectives, what they were, if they were well developed, if they had been achieved and so on, the conclusions. (C1)

> Yes, disparity in the application of criteria. The criteria are established, there is a guide of the FYP where it is stated how it will be evaluated; what happened is that they said that those criteria were not applied or were applied very poorly. Many situations arose that were unfair, and, in addition, there was no precision in the assessment. (P1)

The data show that there are many problems regarding the assessment criteria used by teachers to assess FYP: (a) great disparity in the criteria used; (b) lack of objectivity in the criteria set out in the guide; (c) many teachers do not follow the criteria set out in the guide, but their own. All these problems, together with those mentioned in the previous category, mean that there is a general unease among students and teaching staff with regard to the FYP assessment and grading processes in the faculty, in the sense that there are comparative grievances between the different FYPs and a feeling of inequality and a certain injustice, in the sense that the final grade for an FYP depends in part on who the lecturers on the panel are, and not only on the actual quality of the work.

All these problems have led many of the teachers to consider the need to design a specific assessment instrument, in which the criteria are much more clearly defined and agreed, in order to seek greater reliability in the assessments and marks that each FYP deserves.

Process of creation of the instruments for the assessment of the FYP (years 2015-2017).

In 2015, a group of lecturers put forward the idea of creating some instruments to assess the FYP in a more concrete and precise way. For the analysis of results, the data are divided into three subcategories: (a) motivation to create the instruments and types; (b) process of creation of the instruments; and (c) final process of experimentation.

Motivation to create the instruments and types

Most of the interviewees agreed that they needed a homogeneous and common assessment instrument that they could use in the evaluation of the FYP and marking, although it is true that there were different positions, and even some of them opposed to each other. They also wanted a more objective instrument, which would eliminate discrepancies between teachers, with detailed and common assessment criteria:

> We needed an assessment instrument, something homogeneous, that would serve us all (...) So that was it, to try to look for something objective, to have an instrument. (C1)

> The interest that can be intuited is to make the fairest and most accurate, and also most convenient assessment for the members of the tribunal, because we set exactly as an agreement what we expect a FYP in Education to be and what the previously agreed quality criteria are (...). It was seen as something more objective, which eliminated more immediately the original problem of discrepancy between commissions. (C2)

Only one of the coordinators highlighted as a motivation the fact of making it easier for students to know how their work would be evaluated. In addition, he underlined one more motivation: to have a common FYP assessment instrument to facilitate the assessment process for new lecturers who join the teaching profession; so that it can serve as a guide, both for tutoring a FYP and for evaluating it as a commission:

> The last reason was, when there were new lecturers joining, they were not used to the process. For example, an associate lecturer who has just joined, (...) does not necessarily know how to grade and how to evaluate. (C3)

> If I am bad at spelling, I have to know that if I continue at this level (D), which is the usual one, I will not be able to pass this, so either I revise the document a lot or... (C3)

A coordinator sees the motivation of having an instrument with which to support his assessment if the student questioned their grade:

> Something that would be useful to us so that in the event that a student asked you for a review we could tell him, look we have based ourselves on this and on this. (...) In case a student comes for review, and we could base ourselves on something objective. (C1)

> Ah, that's true! What I have heard the most, if someone complains you have something to base your arguments on, and if not the other way, how do base your arguments? Do you need evidence to answer the student's complaint? You can't carry out an educational process based on the answer to a student's complaint, which is an administrative act. (P1)

Not all lecturers agreed with the creation of the assessment rubrics for the FYP. This is shown by one interviewee, but it was adapted to the majority:

> C2, for example, was convinced of that (against the use of rubrics). I also talked about it with him and discussed it a lot about the rubrics. But well, since that was the approach of the majority, that was that. (P1)

The results indicate that the lecturers agreed on the need to create an assessment instrument for the FYP. This agreement leads us to consider the main characteristics that the rubrics should have: (a) they should offer objectivity; (b) they should be homogeneous for the

FYPs of the entire Faculty of Education with common assessment and marking criteria. This last point is key to helping all types of teachers, regardless of their category and time in the faculty, to assess within the same line of work.

Also, having an instrument with clear, objective criteria created in common with all the teachers in favor generates greater security when assessing and marking the student, both in their written work and in the oral presentation. This security shows that the assessment is reliable based on consciously and collectively created standards, as teachers assess and mark according to criteria related to competences and/or to what is expected to be achieved after the work has been done, if they make good use of the instruments.

The three assessment instruments are analytical in nature. When creating them, they did not specify the type of instrument they wanted (whether analytical or holistic), they just knew that they needed an assessment instrument as detailed and clear as possible. The analytical rubric, being so detailed and closed, allowed them to break with the disparity of criteria that arose among the faculty when evaluating:

> We were also not conscious of the type of rubric. What we wanted was an instrument that was as clear as possible, as unambiguous as possible, and that gave the maximum detail of each of the elements to be evaluated. (C1)

> It was suggested that there were different types of rubrics, but the truth is that the heterogeneity of the participating lecturers encouraged us to have an analytical rubric. If the precedents had been those of a fairly similar assessment, even among commissions and even very similar criteria, there would have been no need for an analytical rubric. But in this case, since there were discrepancies, it was the best solution. (C2)

The data show that there was no objective intentionality in the creation of the analytical instruments. It was the aforementioned motivation that naturally gave rise to this analytical characteristic. The analytical instruments favor what the teachers demanded: that every teacher assesses in the same way and, moreover, justifies a mark by means of fixed criteria. This analytical nature closes the door to flexibility in the assessment of the FYP, an aspect that can only be taken advantage of with the tutor's report (which is the only instrument that does not have weight in itself in the final mark, but which can influence it). In general, it can be seen how the assessment and marking process of the FYP subject is mostly summative and final, unless the assessment instruments are used formatively by the student's tutor.

Process of creation of the instruments

In 2015, a group of lecturers put forward the idea of creating some instruments to assess the FYP in a more concrete and precise way. This proposal was presented to the entire faculty and about 22-23 lecturers participated. They started with an initial exchange of ideas about what each one expected from the FYP; to begin the process of developing the instruments from the same idea, they needed a common and coherent assessment of the FYP. The interviewees agree that there was very open and full participation by the faculty, especially at the beginning of the process:

> The first step was to see what we wanted to evaluate and that responded directly to the fact that we had to create three different instruments, one to evaluate the process, one to evaluate the document and one to evaluate the presentation, because they are like three well differentiated spaces. (C3)

> In the beginning, we, the people from the committee, were the first ones who were involved; but then it was opened up to all the lecturers, of course, but as there were many of us, there were people who were more involved and people who were less involved. The truth is that the load finally was for us. (C1)

> There was a small coordinating committee of four or five lecturers carrying a little bit of the load and leadership of the meetings; but the truth is that, in general, there was a lot of participation because every time we proposed a meeting, we can talk about 22 or 23 lecturers in a seminar discussing the rubrics (....). The sessions were group discussion sessions of the faculty of the degree and there the drafts were presented and discussed. (C2)

The lecturers started from different organizational criteria to create the instruments. They began by discussing what an FYP had to be like, specifying the elements to be addressed in the project, and considered what types of FYPs were to be carried out, so as to adapt the criteria to the nature of the different projects: literature review, practical implementation, research, etc.

> Yes, this was devised to assess, to grade FYPs of students in the education degree, because they were the ones we were going to grade. So, we had in mind the types of work that could be presented. (C1)

They also pointed out the minimum criteria the papers had to meet to be submitted and presented. They adapted and organized the criteria in a coherent way with the objectives of the FYP, the competences that had to be shown in the work to pass the degree, the people who were going to carry out the work and what could be presented, according to what had been worked on throughout the degree. This internal curricular coherence shows that the lecturers organized the assessment of the FYP in such a way that the students saw the significance and logic of the development of the FYP with what was worked on in the degree and in their future profession:

There were discussions from all points of view: from discussions on the issue of spelling (...) Also on the basis of on what criteria we consider the work unacceptable for presentation and defense. From discussions focused on formal aspects to discussions on substance. Whether the FYP should have a research work character or could have an experimental character or was linked or not to the Practicum, to questions even of the approach itself. (...) Yes, it was always contextualized in the FYP of the education degrees. (...) The FYP is (...) evidence of competence. (...) I am able to present a project or work that effectively reflects that it has acquired all those competencies of the degree. (C2)

> Each of these elements correspond to the competencies we are assessing. (C3)

Regarding the report of the FYP elaboration process, the criteria were more flexible because that report was not graded:

> The spirit is the same in all three, but the criteria used and the reasons why they are structured are different. In the one on process, since there was no grading, it allowed us to be more flexible in deciding what to evaluate. We decided in the end that it should be the criteria of the competencies that we are expecting a graduate to be able to perform. (C3)

In all three instruments they defined some basic aspects that would imply failing the FYP, for example. plagiarism and bad spelling. This was a complex aspect to achieve as they had to clearly define each descriptor in order to group all possible projects together; since in this faculty three types of FYP can be presented: (a) theoretical-research; (b) practical (implementation of school curriculum); (c) a combination of theoretical research with practical implementation.

> It was piecemeal. I don't remember having a starting instrument. Because we checked and there was nothing. In fact, afterwards we have seen that people have copied us (...). The FYP is something new in the context of the degrees, so, based on other working rubrics, everyone contributed (...). Then we tried to adjust it to the characteristics of the PJF. (C1)

> Also, based on what criteria we considered the project to be unacceptable for presentation and defense. Discussions focused on formal aspects to substantive discussions. (C2)

> However, we said that there were things that were inadmissible, even if there was only one thing of these, only one aspect at this level, it means that the work is not admissible. (C3)

Once they had the final assessment criteria, they organized them into four qualitative indicators of achievement (A-B-C-D) and then weighted them for the rubrics to make the leap to grading; in each indicator there was a detailed description of the level of achievement in that indicator. There they decided that the evaluation committee was in charge of grading the final paper (80% of the grade: it is divided into indicators, with the first eight indicators of the rubric worth up to 4 points out of 10, and the other three indicators worth up to 2 points, and the other three indicators worth up to 2, making a total of 8 points out of 10) and the oral presentation (20% of the grade), and that the tutor did not grade the final paper, so his or her report had no quantitative value:

> What we had to be clear about was that a priori (valid for all three types of FYP), we should not discard one type of FYP over another, nor discard one over another; that complicated the definition of the items a bit. Part of the complexity was to define items that grouped the different approaches to the papers. (C2)

> So that is associated with the marking (...) If we want to mark, we stick to the percentages that we had established as criteria for distributing the weighting of the mark. So, it fitted perfectly, and that meant that the first one, the process, had no quantitative value in the final mark. (C3)

The three instruments were created with the same purpose, but each with different organizational criteria. When elaborating the rubrics, the lecturers took the following steps: first they specified the aspects to be evaluated, then they set the indicators for each criterion and, finally, they marked the graduation of each indicator. The rubric for the oral presentation was simpler than the two previous ones because it had less weight in the grading. They included in it items that a lecturer should check when speaking in public on an educational topic:

> First it has a very standard structure of what a rubric is, OK, there are the broad aspects, criteria that are set to evaluate and then the levels of achievement and the descriptors or indicators for each cell of the rubric. But this is the final point, of course, previously there were several drafts to reach a final agreement on the rubric. (C2)

> The third rubric was simpler, because it seems that when something has less weight in the rating it relaxes the requirements a little bit. In this case, we were establishing the aspects that a lecturer should be able to control. (C3)

By specifying the nature of each instrument, the tutor's report provides access to information that the committee cannot evaluate by reading the final paper and listening to the oral presentation alone. One of the coordinators states that this report is the least relevant for the evaluating commission:

> In the case of the tutor, well, a rubric was created, but perhaps of the three that there are it is the least relevant, because the tutor's report was to access aspects that the commission cannot see. The committee only sees the result, a product and the defense. (C2)

Based on the results obtained, the importance of collaborative work between professors within the faculties can be seen. They had a handicap, which was the lack of homogeneity in the assessment of the FYP, and they joined forces to solve it and create consensual solutions. This consensus is based on analyzing different aspects of the subject in question: (a) what they expect from a FYP; (b) what aspects in particular should not be missing in the project; (c) what types of FYP there were, in order to include the different types in the instrument; (d) what basic minimums the project had to have in order to be submitted and be able to defend it; (e) what competences should be assessed in the projects, which responded to the degree they were to acquire later; (f) significance in the assessment-criteria-objectives of the FYP within the Degree; (g) what elements implied a failure in the subject. These analyzed indications entail an important detailed study work, where they obtained that a descriptive scale is the instrument that best suits the objectivity and detail, they were looking for in the assessment of the FYP.

Based on the results obtained and the information given by the participants,, in general, the process of creating the instruments followed the following steps: (1) analysis and debate of the problem within the context, concluding that three instruments were needed: to assess the process; to assess the document and to assess the oral presentation and defense; (b) study of the subject of FYP as a teacher's vision and, within it, as a future student (school teachers); (c) application of solutions by groups that could solve the problems found; (d) search for the perfect instrument that would respond to these solutions; © development of evaluation criteria; (f) percentageing of the criteria, according to the graduation of their indicator and general importance to fulfil the objectives of the subject.

Final process of experimentation

The three instruments were applied in the assessment of the FYP of the 2015-2016 academic year. Afterwards, the commission considered that they should modify aspects of that

draft that were not quite clear (writing aspects). From there, the now official and definitive assessment instruments were created; applied from the 2016-2017 course:

> With that draft we did yes, we applied it that year (2015-2016). I do not remember well, it sounds to me something like it did not have a direct link with the qualification, but it was like a kind of support tool. We applied it experimentally in the early call (...) for those who wanted to take the competitive examinations. And since it started well, we began to apply it to the others. The following year things were modified, because with the use of the instrument there were things that were not clear (...). And I do not know if in the second or third year it was no longer reviewed. (C3)

> They stopped reviewing in the 2016-2017 year. With the experience of the first year (2015-2016), adjustments were made. (...) In 2016-2017 the rubric was already set. (C2)

The process of creating the rubrics did not remain in the coordination meetings between the lecturers, but a period of time (a course) was devoted to studying how they worked. When applying them in the assessment and marking of the TFGs, lecturers could see that different aspects of the wording of the indicators of achievement were not distinguishable from each other. This is why it is so important to apply the instruments in real practice, to validate them in concrete situations. Even so, it is a very positive aspect that only parts of the wording were clarified for several reasons: (a) it shows that the instrument created was good, valid and applicable to the educational project work; (b) it closes spaces for the variability of opinions and assessments of the same item by several teachers, as the clearer the indicator, the less differences there will be in its assessment.

Discussion

This section has been organised into two main topics in order to respond to the objectives of this study: (a) to review the procedure and criteria used in the assessment of the FYP in a Spanish faculty of education between 2009-2015, when there were no specific instruments for this purpose, and(b) to analyse the reasons and the process of creating the FYP assessment instruments in 2015 in this faculty.

The first objective focuses on reviewing the procedure and criteria used in the assessment of the FYP in a Spanish faculty of education between 2009-2015, when there were no specific instruments for this purpose. The results show that in 2007, the faculty adapted the assessment criteria of the general guide of FYP (RD 1393/2007) to the specific needs of their faculty. As García and Valle (2014) state, universities have the possibility to develop their own assessment criteria of the FYP. However, the established assessment criteria were very general, and each lecturer interpreted them in his or her own way. This generated several problems in the FYP assessment and grading processes: (a) great disparity in the criteria used; (b) lack of objectivity in the criteria set out in the guide; (c) many professors do not follow the criteria set out in the guide, but their own; (d) there were comparative grievances between the different FYPs and a feeling of inequality and some injustice, in the sense that the final grade of a FYP depends partly on who the professors on the panel are, and not only on the actual quality of the project.. Ruiz (2013) states that, if the assessment criteria are not correctly specified, the assessment in the FIP will not favor the certification and the correct acquisition of the learning and competences acquired in the degree.

For all these reasons, the teachers of the faculty should consider the need to design a specific assessment instrument (rubrics), in which the criteria are much more clearly defined and agreed upon, in order to seek greater reliability in the assessments and qualifications that each FYP deserves. The choice to specify the elements of the FYP subject coincides with Biggs (2005) and his approach to curricular alignment. In this way, they would avoid the difficulty in reaching a consensus among the assessments, since they did not have common and concrete criteria; In this respect, Mellado (2013) considers that rubrics are designed to evaluate objectively, for which it is necessary to clearly specify what is expected from the student with the task and what are the criteria considered to evaluate it. This approach coincides with what the results indicate about the assessments of the interviewees and the procedure followed.

There are two rubrics for the assessment and grading of the FYP and its public defense. In this case, the analytical rubrics were the most suitable for an assessment of the FYP in the faculty with criteria as closed as possible, where there was no room for different interpretations among lecturers. According to Ruiz (2015) the rubrics of this work are analytical, as they detail and break down the elements to be evaluated.

The tutor's report is an instrument that can serve for the committee to have a global vision of the elaboration process and the quality of the FYP; so that it does not evaluate the final result in isolation. Vera and Briones (2014) affirm that students positively value the work of tutoring because the tutor is the only one who develops learning with the FYP in the student; therefore, his report is the one that reflects the evolution of the acquired learning. Jonsson (2014) affirms as well that the formative, transparent use of rubrics helps students get to know the assessment criteria in order to self-assess, which, in turn, facilitates student understanding and use of the criteria. However, there are some authors that take on a different point of view (Torrance, 2012). They support the idea that the most important aim of higher education is a student's independent thinking, not convergent thinking caused by transparent assessment criteria and processes. The rubric has no value in itself, what matters is its formative use, which is where gaps and issues are often observed. In general, Panadero and Jonsson (2020) have shown after an extensive literature review that the correct use of rubrics helps students to improve their learning, but it is not clear in which specific sense.

We believe that these results may be transferable to other faculties of education where they do not have detailed assessment instruments but work with general assessment criteria that appear only in the FYP guide. The second research objective was to analyse the reasons for and the process of creating the 2015 FYP assessment instruments in the faculty.

The idea of arriving at more homogeneous and less disparate assessments among teachers seemed to be one of the triggers for the creation of rubrics. As soon as there are different teachers with different opinions and without a concrete instrument, assessment becomes a real seesaw of visions and discussions.

Only one lecturer emphasized that the rubrics could help to guide the FYP from the beginning; coinciding with Biggs' (2005) idea that the assessment should serve as a guide to orient the student in the teaching-learning process, in the sense of the English term "feedforward." Ion, Silva, and García (2013) define this term as "prospective feedback", which helps the learner to learn about the task they are performing so that they will know how to perform it correctly in the future.

There is a motivation in the creation of rubrics that is particularly striking: to justify possible student complaints. This idea does not appear in the works that present the purposes of the rubric as an assessment tool (Fernández-March, 2011; Martínez et al., 2013). Therefore, we see that rubrics are instruments with a clear purpose of summative and qualifying assessment and that this seems to be a widespread assessment system; as shown by García and Valle (2014) in a study on a sample of Spanish universities.

On the other hand, the process of elaboration of the rubrics for the assessment of the FYP was coordinated by three faculty lecturers. In this regard, they comply with what several authors indicate about the advantages of the use of rubrics in the assessment in higher education (Quintana & Gil, 2015) and, specifically, with the authors who defend an assessment of the FYP taking rubrics as the main instrument (Vera & Briones, 2014).

Before entering the technical process of developing the rubrics, the interviewees analyzed and organized the assessment criteria with the objectives and competencies of the FYP, as well as the coherence of the work with their future profession and what they had worked on in undergraduate studies. This approach is developed by Fernández-March (2016), as he considers that the competency-based approach should allow the university lecturer to evaluate the student in terms of his or her ability to put this knowledge into practice to solve problems in real contexts.

For their elaboration, they took as a reference other similar instruments used in the degree and adapted them to the subject. Aspects similar to those proposed by García et al. (2016) were suggested: first they perform an analysis of competencies that the student must show that he/she has acquired in the completion of his/her FYP and, subsequently, they point out indicators and levels of achievement. This elaboration process is similar to that proposed by Mertler (2001) and López-Pastor and Pérez-Pueyo (2017), adding as an initial step to analyze the context of the subject in general and, as a step prior to implementing the instrument, to validate the rubric between the lecturer and the students. All the lecturers gave their opinion in the elaboration process, highlighting the idea of the majority that the so-called "rubrics" were the instruments that best adapted to the ideas put forward. According to López-Pastor and Pérez-Pueyo (2017), "rubrics" are scales with different grades or levels, usually 4 or 5, which contain detailed descriptions of the characteristics that the task being asked for may have. In this respect, they comply with what several authors indicate about the advantages of the use of rubrics in the assessment in higher education (Marín et al., 2012; Martínez et al., 2013) and, specifically, with the authors who defend an assessment of the FYP taking the rubrics as the main instrument (Vera & Briones, 2014).

The three assessment instruments were created in 2015 and were validated during the 2016-2017 academic year. This validation process served to correct certain aspects of wording and concreteness of criteria. The validation of the instruments during an academic year provided an opportunity to improve them and make them even more concrete, as there were different aspects in the achievement indicators that did not seem to be understood. This step was one of the most important, as these instruments were validated in real assessment and marking situations. It could be seen how the rubrics encourage the assessment and grading of FYP, as long as they are used with the same approach, whoever the teacher is assessing. Along the same lines, Fernández-Garcimartín et al. (2022) indicate different strengths of these instruments, including the fact that they facilitate assessment and consensus among the assessment committee.

In essence, the results obtained coincide with different studies published on rubrics as elements of assessment of different learning tasks. Even so, there are no studies that analyze the procedure for creating them. In this sense, the contribution of this study is strong and novel, as it deals with the analysis of needs for the creation of assessment instruments and how this process is carried out (as an example in a specific faculty of education).

This is the reason why, although it is a case study of a unique and particular context, the results obtained are perfectly transferable to different faculties, national and international. Furthermore, the prospective of the subject of the study is broad and is linked to the study of the validity and reliability of these rubrics through a longitudinal study in the faculty itself. We are also considering how to improve and adapt these instruments to the demands of the new, more digitalized university.

Conclusions

The procedure and criteria for the assessment of the FYP used in a faculty of education when there were no specific instruments for this purpose have been reviewed. The criteria were very general and not very objective; hence, each lecturer followed a different and, in many cases, inadequate assessment procedure.

Due to this problem, the faculty decided to create three assessment instruments: a tutor's report and two rubrics for the final assessment of the FYP. An analysis of the process of creating them has been carried out.

The instruments were designed based on the degree competencies and the FYP guide. The process of creating the instruments was carried out taking into account different aspects so that the assessment of the FYP would be adapted to the subject, the competencies and the students. These three instruments were implemented in the 2015-2016 academic year and, after a small validation process, the official instruments were created in the following academic year.

The contribution of this research is the study of a novel topic, which has been scarcely researched at present. There is a considerable lack of studies on rubrics as instruments for the assessment of the FYP in Spanish faculties of education. For this reason, it is of particular relevance and transferability to other faculties, as the FYP is a fundamental element in which the student has to reflect the diversity of competences acquired.

This study makes several important contributions to the scientific and professional community: (a) it studies and shows the need for detailed descriptive scales (rubrics) in order to gain in objectivity and fairness in FYP assessment and grading processes; (b) PYFs are a final subject that must be taken in all undergraduate university studies since the implementation of the EHEA in Spain and other European countries, so the degree of transferability and application to other faculties and contexts is very large, both in Teacher Education and in the rest of studies; (c) the development and validation of these instruments has generated a considerable improvement when it comes to carrying out fairer and more equitable assessment and grading processes in the faculty, so the research has a positive transfer towards improving the social reality.

For all these reasons, we believe that this study may be of great interest to teachers in charge of Teacher Education, as well as to researchers specializing in the field.

A possible limitation of this study is the low sample of subjects interviewed. As a perspective, it could be interesting to analyze the advantages and disadvantages of the application of these assessment instruments over the last eight academic years, and to investigate in which other Spanish universities are using these or similar instruments, and with what results, with a larger sample and also incorporating the students' perspective. Another possibility is to analyze the possibility of carrying out a new process of adaptation and improvement of the instruments on the basis of this previous assessment.

References

Álvarez-Álvarez, C., & San Fabián-Maroto, J. L. (2015). La elección del estudio de caso en investigación educativa. *Gazeta de Antropología, 28*(1).

Biggs, J. (2005). *Quality of university learning*. Narcea.

Black, P., & William, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Assessment and Accountability*, *21*(1), 5-31. https://doi.org/10.1007/s11092-008-9068-5

Brookhart, S. M. (2018). Appropriate criteria: Key to effective rubrics. *Frontiers in Education, 3*(22), 1-12. https://doi.org/10.3389/feduc.2018.00022

Brooks, C., Carroll, A., Gillies, R. M., & Hattie, J. (2019). A matrix of feedback for learning. *Australian Journal of Teacher Education, 44*(4), 2. https://doi.org/10.14221/ajte.2018v44n4.2

Ceballos-Herrera, F. (2009). El informe de investigación con estudio de casos. *Magis, Revista Internacional de Investigación en Educación*, 2, 413-423.

Creswell, J. W. (2013). *Qualitative inquiry and research design* (3rd ed.). Sage. https://doi.org/10.1177/1524839915580941

De Sande, J. C., Eckert, M., Gutierrez-Arriola, J. M., Pescador, F., Garcia-del-Pino, N., Sáenz, L. & Fraile, R. (2011). Competence assessment of final year project for undergraduate telecommunication students. In *4th International Conference of Education, Research and Innovation*, 989–997. IATED.

Fernández-Garcimartín, C., López-Pastor, V. M., Fuentes Nieto, T., & Hortigüela-Alcalá, D. (2022). Los instrumentos de evaluación de los trabajos Fin de grado. *Revista INTEREDU, 1*(6), 97-126. https://doi.org/10.32735/S2735-652320220006114

Fernández-March, A. (2011). Learning-oriented assessment in a competency-based training model in university education. *REDU, 8*(1), 11-34. https://doi.org/10.4995/redu.2010.6216

García, M. A., & Valle, A. M. (2014). Comparative analysis of the regulations on the final thesis of the mathematics degree in state universities. In Pérez et al. (Coords.). *Proceedings of the I Interuniversity Conference on the End-of-Degree Project* (pp. 149-158). Bilbao.

Guba, E. G. (1989). *Credibility criteria in naturalistic research. Teaching: its theory and practice.* Akal.

Hashim, N., & Hashim, H. (2010). Outcome based education performance assessment on the Final Year Degree Project. In *The Proceedings of the 7th WSEAS international conference on engineering education*, 215–222. https://doi.org/10.5555/1864181.1864223

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112. https://doi.org/10.3102/003465430298487

Huaman, L. A. (2015). Rubrics as an assessment instrument in university students of the Faculty of Education-UNCP. *Prospectiva Universitaria, 12*(1), 93-104. https://doi.org/10.26490/uncp.prospectivauniversitaria.2015.12.460

Ion, G., Silva, P., & Cano García, E. (2013). El feedback y el feedforward en la evaluación de las competencias de estudiantes universitarios. *Profesorado, 17*(2), 283-301. http://hdl.handle.net/10481/30020

Jawitz, J., Shay, S., & Moore, R. (2002). Management and assessment of final year projects in engineering. *International Journal of Engineering Education*, *18*(4), 472-478. https://cutt.ly/ZjkntN6

Jonsson, A. (2014). Rubrics as a way of providing transparency in assessment. *Assessment and Assessment in Higher Education, 39*(7), 840-852. https://doi.org/10.1080/02602938.2013.875117

Krause, M. (1995). La investigación cualitativa: un campo de posibilidades y desafíos. *Revista temas de educación, 7*(7), 19-40. Recuperado de https://cutt.ly/DgDNojr

López-Pastor, V. M., & Pérez-Pueyo, Á. (2017). *Evaluación formativa y compartida en educación: experiencias de éxito en todas las etapas educativas* (e-book). Universidad de León. https://buleria.unileon.es/handle/10612/5999

Marín, M., Cabero, J., & Barroso, J. (2012). The assessment rubric in the university teacher training process. The DIPRO project proposal 2.0. *Educar, 48*(2), 347-364. https://doi.org/10.5565/rev/educar.29

Martínez, E., Tellado, F. & Raposo, M. (2013). The rubric as an instrument for self-assessment:

a pilot study. *REDU, 11*(2), 373-390. https://doi.org/10.4995/redu.2013.5581

Mateo, J., Escofet, A., Martínez, F., Ventura, J. J., & Vlachopoulos, D. (2012). The Final Year Project (FYP) in social sciences: Establishment of its associated competences and assessment standards. *Studies in educational assessment, 38*(1), 28-34. https://doi.org/10.1016/j.stueduc.2011.12.002

Maxwell, J. A. (2019). *Diseño de investigación cualitativa. Un enfoque interactivo*. GEDISA.

Medina, C., Begoña De la Iglesia, S. V., & and Roselló, M. R. (2020). Design, application and assessment of the formative feedback for FYP tutoring. *Magister*, *32*(1), 1-8. https://doi.org/10.17811/msg.32.1.2020.1-8

Mellado, E. (2013) *Development and assessment of competencies in pre-service teacher education using e-portfolio and rubric*. [Doctoral Thesis, University of Extremadura]. https://cutt.ly/Og8ATOK

Mertler, A. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research, and Assessment, 7*(1), 25. https://doi.org/10.7275/gcy8-0w24

Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, *9*, 129-144. https://doi.org/10.1016/j.edurev.2013.01.002

Panadero, E., & Jonsson, A. (2020). A critical review of the arguments against the use of rubrics. *Educational Research Review, 30,* 100-329. https://doi.org/10.1016/j.edurev.2020.100329

Quintana, M., & Gil, J. E. (2015) Rubrics as an effective method of assessment in the assessment of learning. *Alternativas, 16*(3), 5-13. https://doi.org/10.23878/alternativas.v16i3.73

Royal Decree 1393/2007, of October 29, 2007, which establishes the organization of official university education (BOE no. 260, of 10/30/2007). https://cutt.ly/Cg8AyS4

Ruiz, R. (2015). *Análisis de las rúbricas para la evaluación de la expresión oral y escrita en estudiantes de grado de magisterio en educación primaria*. [Doctoral thesis, UNED]. https://cutt.ly/3g8Af4r

Sancho-Esper, F. M. (coord.) (2018). Estrategias de mejora en la elaboración y evaluación de los FYPs: la rúbrica como herramienta de auditoría. In Roig-Vila, R. (Coord..) *Memorias del Programa de Redes-ICE de calidad, innovación e investigación en docencia universitaria. Convocatoria 2017-2018* (pp. 185-204). Alicante: Universidad de Alicante.

Sharef, N., Hamdan, H., & Madzin, H. (2014). Innovation-enhanced rubrics assessment for final year projects. *Global Journal of Engineering Education*, *16*(3), 129-135. http://www.wiete.com.au/journals/GJEE/Publish/vol16no3/05-Sharef-N.pdf

Stake, R. E. (1995). *The art of case study research.* Sage Publications.

Stake, R. E. (2005). *Investigación con estudio de casos*. Morata

Torrance, H. (2012). Formative assessment at the crossroads: Conformative, deformative and transformative assessment. *Oxford Review of Education, 38*, 323-342. https://doi.org/10.1080/03054985.2012.689693

Vaismoradi, M., Jones, J., Turunen, H., & Snelgrove, S. (2016). Theme development in qualitative content analysis and thematic analysis. *Journal of Nursing Education and Practice, 6*(5), 100–110. https://doi.org/10.5430/jnep.v6n5p100

Vera, J. & Briones, E. (2014). Elementos clave del proceso de tutorización y evaluación del FYP: análisis y resultados en la UVA y la UC. In K. Pérez, B. Bilbao, E. Fernández de Larrea, B. Molero, & P. Ruiz de Gauna (Eds.), *Actas del I Congreso Interuniversitario sobre el Trabajo de Fin de Grado. Retos y oportunidades del FYP en la sociedad del conocimiento* (pp. 335–345). Universidad del País Vasco.

Vicario-Molina, I., Martín, E., Gómez, A., & González, L. (2020). Nuevos desafíos en la

Educación Superior: análisis de resultados obtenidos y dificultades experimentadas en la realización del Trabajo de Fin de Grado de estudiantes de los Grados de Maestro de la Universidad de Salamanca. *Revista complutense de educación, 31*(2), 195-194. https://doi.org/10.5209/rced.62003

Zhang, Y., & Wildemuth, B. M. (2009). Qualitative analysis of content. In B. Wildemuth (Ed.), *Applications of social research methods to questions in information and library science* (pp. 2-12). Libraries Unlimited.

Author Note

Carla Fernández Garcimartín, PhD in Transdisciplinary Research in Education (University of Valladolid, FPU contract) and Master in Research in Social Sciences, Education, Audiovisual Communication, Economics and Business by the University of Valladolid, is a Graduate as a teacher in Primary Education (Physical Education) and Early Childhood Education (University of Valladolid). She is Member of the International Network of Formative and Shared Assessment in Education. Her main lines of research: physical education, formative and shared assessment, initial teacher training, Final Degree Projects. Please direct correspondence to carlafdzg96@gmail.com

Víctor Manuel López Pastor, PhD in Education since 1998, is University Professor and full-time professor at the Faculty of Education of Segovia, Coordinator of the International Network of Formative and Shared Assessment in Education, and Director of 17 doctoral theses and author of numerous books and scientific articles. His main lines of research: formative and shared assessment, initial and continuing teacher training, action research and teaching innovation, didactics of physical education and early childhood education. Please direct correspondence to victor.lopez.pastor@uva.es

Teresa Fuentes Nieto, PhD in Physical Activity and Sport Sciences, extraordinary doctorate award 2011 (University of Las Palmas) and Degree in Physical Activity and Sport Sciences (Universidad Politécnica de Madrid), is Associate Professor at the Faculty of Education of Segovia, Teacher of Physical Education in Secondary Education (IES Maria Moliner), and Member of the International Network of Formative and Shared Assessment in Education. Her main lines of research are physical education, formative and shared assessment, initial teacher training, in-service teacher training and Final Degree Projects. Please direct correspondence to fuentes.teresa@gmail.com

David Hortigüela Alcalá, International Doctor in Education (Excellent Cum Laude and extraordinary award), is Professor and Director of the Area of Didactics of Corporal Expression. Director of the Department of Specific Didactics. He is a Specialist in pedagogical models and open and participatory methodologies in the area of Physical Education, in formative and shared assessment, in research-action work cycles in the classroom and in pre-service teacher education. He is author of more than 300 publications on the educational field and speaker at various conferences, courses and seminars related to teacher training. He has carried out research stays in more than 10 countries. Please direct correspondence to dhortiguela@ubu.es

Acknowledgements: This research was supported by Grant RTI2018-093292-B-I00 funded by MCIN/AEI/ 10.13039/501100011033 and by "ERDF A way of making Europe."

Article Citation

Fernández-Garcimartín, C., López-Pastor, V. M., Fuentes Nieto, T., Hortigüela-Alcalá, D. (2024). Creating assessment rubrics for final teacher education degree projects: A qualitative case study. *The Qualitative Report, 29*(6), 1689-1711. https://doi.org/10.46743/2160-3715/2024.6426