

COMPARATIVE RESULTS OF PEER ASSESSMENT vs EVALUATION SUPPORTED IN RUBRICS BOTH IN UNDERGRADUATE AND MASTER'S DEGREE STUDIES

N. Montanes^{1*}, L. Quiles-Carrillo², D. Garcia-Sanoguera³, V. Fombuena⁴, O. Fenollar⁵

¹Prof. Dr. Grupo de Innovación de Prácticas Académicas (GIPA), Universitat Politècnica de València (UPV), SPAIN, nesmonmu@upvnet.upv.es

²Prof. Mr. Grupo de Innovación de Prácticas Académicas (GIPA), Universitat Politècnica de València (UPV), SPAIN, luiquic1@epsa.upv.es

³Prof. Dr. Grupo de Innovación de Prácticas Académicas (GIPA), Universitat Politècnica de València (UPV), SPAIN, dagarsa@dimm.upv.es

⁴Prof. Dr. Grupo de Innovación de Prácticas Académicas (GIPA), Universitat Politècnica de València (UPV), SPAIN, vifombor@upv.es

⁵Prof. Dr. Grupo de Innovación de Prácticas Académicas (GIPA), Universitat Politècnica de València (UPV), SPAIN, ocfegi@epsa.upv.es

*Corresponding author

Abstract

The process to evaluate students is always a committed task. Evaluation can be a way of regulating the student lessons and also a tool to help them to improve. In order to achieve our goal, teaching, learning and evaluation should be aligned, which requires clear criteria and indicators. The use of rubrics as an evaluation tool allows the establishment of impartial and clear indicators. In addition, the rubrics promote formative evaluation, which is very interesting to treat evaluation as one more formative activity.

In the described framework of a learning-oriented evaluation, peer assessment, is to be evaluated by one's own peers. The process encourages active learning in the students and become the protagonists of their own learning process.

With the aim of achieving this formative evaluation, the authors have established peer-to-peer assessment in certain subjects of the different Degrees in Engineering, as well as in the Master's Degree in Engineering, Processing and Characterization of Materials, both taught at the Universitat Politècnica de València (UPV). Specifically, rubrics have been developed to evaluate the oral presentations that students make of their projects, and the same rubric is used both by students for peer evaluation and by the lecturer.

The aim of this work is to show the results registered after several years of application of the described evaluative methodology and to compare them in order to obtain conclusions which validate their application in the different formative levels in the field of the universities.

Keywords: Peer assessment, evaluation, rubrics.

1 INTRODUCTION

Establishing an appropriate evaluation system, as well as the act of evaluating itself, is always a committed task for teachers. Evaluation can be a way of regulating what the student learns and also a tool to help them to improve. One way to proceed with the evaluation is to use the constructive alignment proposed by John Biggs. Teaching is reinforced by aligning hers objectives, hers methods and the assessment tasks (Biggs & Biggs, 2004). For this, clear criteria and indicators are needed.

And it is that even the evaluation protocols for the verification of official university degrees (Agencia Nacional de Evaluación de la Calidad y Acreditación - ANECA), and for the renewal of accreditation (Programa ACREDITA de la ANECA), in Spain, establish as a guideline that there is coherence or alignment between learning results and formation and evaluation activities (Fig. 1).

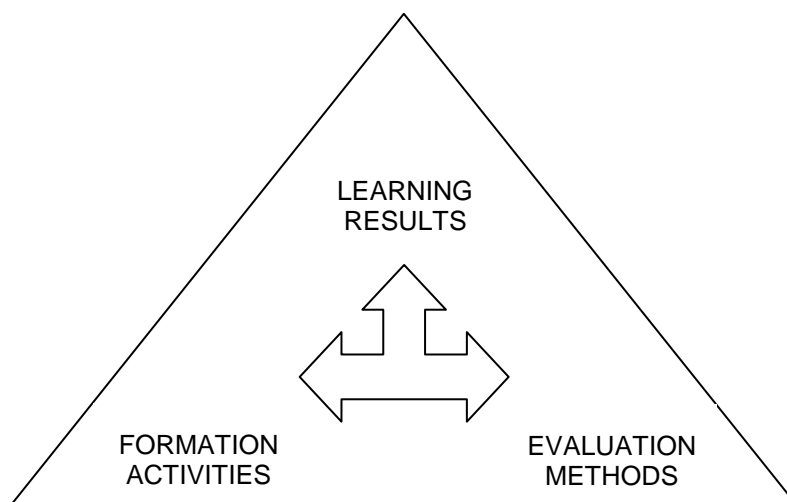


Fig. 1. Triangulation between Formation Activities, Learning Results and Evaluation Methods

The triangulation of these three concepts it is of vital importance to guarantee the quality of teaching and to reinforce the focus of the student-centered teaching and learning process. It should be remembered that there is a convergence of the educational systems of European countries towards a European Higher Education Area (EHEA), being the central axis of the change to make the student protagonist and responsible for their own learning, which must be meaningful and autonomous (Sánchez & Ruiz, 2011).

If the learning results are ambiguous, difficult to understand and/or complicated to achieve throughout the course, the evaluation will be complex. In this regard, the use of rubrics allows the establishment of clear and impartial indicators. As Stevens and Levis proposed, "a rubric is a scoring tool that lays out the specific expectations for an assignment. Rubrics divide an assignment into its component parts and provide a detailed description of what constitutes acceptable or unacceptable levels of performance for each of those parts" (Stevens & Levi, 2013). The rubrics used as an evaluation tool enhance the formative evaluation, being that it is very interesting to treat the evaluation as one more didactic activity.

In the described framework of a learning-oriented evaluation, peer assessment, is to be evaluated by one's own peers, encourages active learning in which the students also become the protagonists of their own learning process (Sluijsmans, Brand-Gruwel, & van Merriënboer, 2002). Through this system, students compare their work, their efforts, their results, with those of their peers in the act of evaluation, and this comparison produces a feedback that drives students towards improvement.

With the aim of achieving this formative evaluation, the authors have begun to implement in certain subjects an evaluation methodology in which the evaluation of the teacher is combined with the peer assessment, using the same rubric. For this moment, the evaluative methodology described is being applied only to evaluate the oral presentations that the students make of their projects, but the objective is to apply it progressively in other acts of evaluation of the subjects.

The aim of this work is to show the results registered after several years of application of the described evaluative methodology and to compare them in order to obtain conclusions which validate their application in the different formative levels in the field of the universities.

2 METHODOLOGY

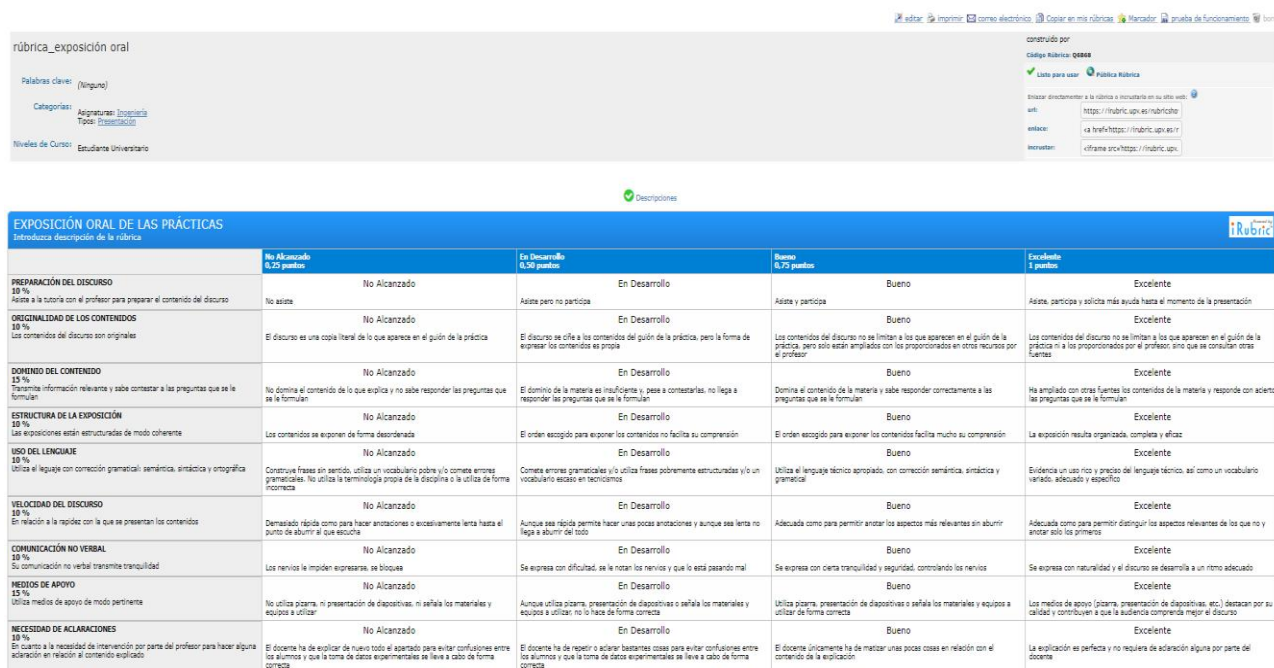
In certain subjects of the different Degrees in Engineering, as well as in the Master's Degree in Engineering, Processing and Characterization of Materials, both taught at the Universitat Politècnica de València (UPV), a project-based learning is carried out. The results of their project must be defended by the students at the end of the subject through an oral presentation.

At the beginning, these final presentations of the projects were evaluated only by the lecturer. However, in order to achieve a significant improvement in their teaching, the authors have started to apply other techniques in which the evaluation system itself promotes learning, including peer assessment.

The score obtained in the final presentation - defense of a project, of a student, is the average of the score given by the classmates and the score given by the teacher. In order to have students and teacher a common criterion in the evaluation, clear and precise indicators are established, describing in them the different levels of development reached and their corresponding weighting. That is, a rubric has been used as an evaluation tool.

It is worth mentioning that the UPV has an on-line platform to support on-site teaching called "PoliformaT". Among other utilities, in this platform is the tool "irubrics" in which, among other things, it is possible to consult UPV institutional rubrics created for different purposes. From this repository of rubrics, the lecturers involved in this work chose one to evaluate the oral expositions of the engineering students, which is shown in the Fig. 2 and which can be accessed at the address:

<https://irubric.upv.es/rubricshowc.cfm?code=Q6B68&sp=yes>



| | No Alcanzado 0,00 puntos | No Alcanzado | En Desarrollo 0,50 puntos | En Desarrollo | Bueno 1,00 puntos | Bueno | Excelente 1 puntos | Excelente |
|--|---|--------------|--|---------------|---|-------|--|-----------|
| PREPARACIÓN DEL DISCURSO 10% Ayuda a la tutoría con el profesor para preparar el contenido del discurso | No asiste | No Alcanzado | Asiste pero no participa | En Desarrollo | Asiste y participa | Bueno | Asiste, participa y solicita más ayuda hasta el momento de la presentación | Excelente |
| ORIGINALIDAD DE LOS CONTENIDOS 10% Los contenidos del discurso son originales | El discurso es una copia literal de lo que aparece en el guión de la práctica | No Alcanzado | El discurso se cifra a los contenidos del guión de la práctica, pero la forma de expresar los contenidos es propia | En Desarrollo | Los contenidos del discurso no se limitan a los que aparecen en el guión de la práctica, pero sí están ampliados con los proporcionados en otros recursos por el profesor | Bueno | Los contenidos del discurso no se limitan a los que aparecen en el guión de la práctica ni a los proporcionados por el profesor, sino que se consultan otras fuentes | Excelente |
| DOMINIO DEL CONTENIDO 15% Transmite información relevante y sabe contestar a las preguntas que se le formulan | No domina el contenido de lo que explica y no sabe responder las preguntas que se le formulan | No Alcanzado | El dominio de la materia es insuficiente y, pese a contestarlas, no llega a responder las preguntas que se le formulan | En Desarrollo | Domina el contenido de la materia y sabe responder correctamente a las preguntas que se le formulan | Bueno | Ha ampliado con otras fuentes los contenidos de la materia y responde con acierto las preguntas que se le formulan | Excelente |
| ESTRUCTURA DE LA EXPOSICIÓN 10% Las exposiciones están estructuradas de modo coherente | Los contenidos se exponen de forma desordenada | No Alcanzado | El orden escogido para exponer los contenidos no facilita su comprensión | En Desarrollo | El orden escogido para exponer los contenidos facilita mucho su comprensión | Bueno | La exposición resulta organizada, completa y eficaz | Excelente |
| USO DEL LENGUAJE 10% Utiliza el lenguaje con corrección gramatical, semántica, sintáctica y ortográfica | Construye frases sin sentido, utiliza un vocabulario pobre y/o comete errores gramaticales. No utiliza la terminología propia de la disciplina o le utiliza de forma incorrecta | No Alcanzado | Comete errores gramaticales y/o utiliza frases pobremente estructuradas y/o un vocabulario escaso en tecnicismos | En Desarrollo | Utiliza el lenguaje técnico apropiado, con corrección semántica, sintáctica y gramatical | Bueno | Evidencia un uso rico y preciso del lenguaje técnico, así como un vocabulario variado, adecuado y específico | Excelente |
| VELOCIDAD DEL DISCURSO 10% En relación a la rapidez con la que se presentan los contenidos | Demasiado rígida como para hacer anotaciones o escucha lenta hasta el punto de aburrir al que escucha | No Alcanzado | Aunque sea rígida permite hacer unas pocas anotaciones y aunque sea lenta no llega a aburrir del todo | En Desarrollo | Adecuada como para permitir anotar los aspectos más relevantes sin aburrir | Bueno | Adecuada como para permitir distinguir los aspectos relevantes de los que no y anotar sólo los primeros | Excelente |
| COMUNICACIÓN NO VERBAL 10% Su comunicación no verbal transmite tranquilidad | Los nervios le impiden expresarse, se bloquea | No Alcanzado | Se expresa con dificultad, se le notan los nervios y que lo está pasando mal | En Desarrollo | Se expresa con cierta tranquilidad y seguridad, controlando los nervios | Bueno | Se expresa con naturalidad y el discurso se desarrolla a un ritmo adecuado | Excelente |
| MEDIOS DE APOYO 15% Utiliza medios de apoyo de modo pertinente | No utiliza pizarra, ni presentación de diapositivas, ni señala los materiales y equipos a utilizar | No Alcanzado | Aunque utiliza pizarra, presentación de diapositivas o señala los materiales y equipos a utilizar, no lo hace de forma correcta | En Desarrollo | Utiliza pizarra, presentación de diapositivas o señala los materiales y equipos a utilizar de forma correcta | Bueno | Los medios de apoyo (pizarra, presentación de diapositivas, etc.) destacan por su calidad y contribuyen a que la audiencia comprenda mejor el discurso | Excelente |
| NECESIDAD DE ACLARACIONES 10% En cuanto a la necesidad de intervención por parte del profesor para hacer alguna aclaración en relación al contenido explicado | El docente ha de explicar de nuevo todo el apartado para evitar confusiones entre los alumnos y que la toma de datos experimentales se lleve a cabo de forma correcta | No Alcanzado | El docente ha de repetir o aclarar bastantes cosas para evitar confusiones entre los alumnos y que la toma de datos experimentales se lleve a cabo de forma correcta | En Desarrollo | El docente únicamente ha de matizar unas pocas cosas en relación con el contenido de la explicación | Bueno | La explicación es perfecta y no requiere de aclaración alguna por parte del docente | Excelente |

Fig. 2. Rubric to evaluate oral expositions of university engineering students

The indicators to be evaluated considered in the rubric are the following:

- Preparation of the speech (10%).
- Originality of the contents (10%).
- Domain of content (15%).
- Structure of the presentation (10%).
- Use of language (10%).
- Speech speed (10%).
- Non-verbal communication (10%)

- Support means (15%).
- Need for clarification (10%).

Finally, Table 1 shows relevant information in relation to the subjects in which the evaluation methodology described has been applied.

Table 1. Set of subjects in which peer assessment / teacher evaluation has been applied through rubrics

| Subject | Degree/Master | Course | Year | Nº Students |
|--|--|--------|-----------|-------------|
| Experimentation in chemical engineering I | Degree in Chemical Engineering | 2º | 2017/2018 | 27 |
| Chemical engineering Bases | Degree in Chemical Engineering | 2º | 2018/2019 | 26 |
| Packaging | Degree in Industrial Design Engineering and Product Development | 3º | 2016/2017 | 54 |
| Packaging | Degree in Industrial Design Engineering and Product Development | 3º | 2017/2018 | 72 |
| Experimentation in chemical engineering II | Degree in Chemical Engineering | 3º | 2018/2019 | 32 |
| Materials, design and restyling | Degree in Mechanical Engineering | 4º | 2018/2019 | 26 |
| Diagnosis and behavior in service | Master's Degree in Engineering, Processing and Characterization of Materials | 2º | 2016/2017 | 9 |
| Diagnosis and behavior in service | Master's Degree in Engineering, Processing and Characterization of Materials | 2º | 2017/2018 | 8 |
| Diagnosis and behavior in service | Master's Degree in Engineering, Processing and Characterization of Materials | 2º | 2018/2019 | 9 |

3 RESULTS

For the set of subjects described in the previous section, Table 2 shows the average results of the peer assessment and the average results of the evaluation of the teacher responsible for the subject, as well as the difference registered between both evaluations. It should be noted that in these subjects the final presentation's score obtained by the students, as stated above, is the average of the score given by the classmates and the score given by the teacher.

Table 2. Results of peer assessment / teacher evaluation

| Subject | Course | Year | Average result peer assessment | Average result teacher evaluation | Difference |
|--|--------|-----------|--------------------------------|-----------------------------------|------------|
| Experimentation in chemical engineering I | 2º | 2017/2018 | 6.54 | 6.11 | 0.43 |
| Bases in chemical engineering | 2º | 2018/2019 | 6.95 | 6.09 | 0.86 |
| Packaging | 3º | 2016/2017 | 8.73 | 8.07 | 0.66 |
| Packaging | 3º | 2017/2018 | 8.21 | 8.29 | -0.08 |
| Experimentation in chemical engineering II | 3º | 2018/2019 | 7.12 | 6.45 | 0.67 |
| Materials, design and restyling | 4º | 2018/2019 | 8.17 | 8.45 | -0.28 |
| Diagnosis and behavior in service | 2º | 2016/2017 | 7.71 | 7.92 | -0.21 |
| Diagnosis and behavior in service | 2º | 2017/2018 | 7.21 | 8.25 | -1.04 |
| Diagnosis and behavior in service | 2º | 2018/2019 | 8.73 | 9.58 | -0.85 |

From the analysis of the results obtained, the following partial conclusions can be drawn:

- The undergraduate students of the lowest courses (2nd and 3rd) normally give their classmates a score slightly higher than the teacher in average value.
- As they near the end of their undergraduate studies (4th year), the teacher's score is somewhat higher than those they give themselves in peer assessment.
- This trend is more accentuated in master's student. The teacher gives higher scores than those given among the students, probably due to the incipient competitiveness created.
- This is probably due to the fact that the students of the last courses are more aware of the effort that the work has cost them and are less lax when it comes to evaluating the result of the work of others compared to their own.
- In any case, the difference between peer assessment and teacher evaluation is below 1 point out of 10, except for some exceptions, which indicate the correct evaluations process used.

4 CONCLUSIONS

In view of the results obtained and the partial conclusions presented in the previous section, the evaluation methodology described seems to provide good results, since the differences between the peer assessment and the teacher evaluation are not great and instead with the evaluation between peers a learning-oriented assessment is achieved, which is highly beneficial for students.

Although one should expect to have more data to draw more firm conclusions, it seems that the methodology described is apt to be applied in more evaluation acts within the subjects presented (or in others), as well as at different formative levels in the field of the universities.

5 ACKNOWLEDGEMENT

This work is the result of the effort being made by the *Grupo de Innovación de Prácticas Académicas (GIPA)* of the UPV in favor of teacher improvement. The authors want to show their gratitude to the UPV and the

Instituto de Ciencias de la Educación (ICE) of the UPV for their help and collaboration in the formation of the *Equipo de Innovación y Calidad Educativa (EICE)* called GIPA.

This work is supported by the *Vicerrectorado de Estudios, Calidad y Acreditación* of the *Universitat Politècnica de València* under PIME program (Ref: B18/18).

REFERENCE LIST

- Biggs, J., & Biggs, J. B. (2004). *Calidad del aprendizaje universitario* (Vol. 7): Narcea ediciones.
- Sánchez, A. V., & Ruiz, M. P. (2011). Evaluación de competencias genéricas: principios, oportunidades y limitaciones. *Bordón. Revista de pedagogía*, 63(1), 147-170.
- Sluijsmans, D. M. A., Brand-Gruwel, S., & van Merriënboer, J. J. G. (2002). Peer Assessment Training in Teacher Education: Effects on performance and perceptions. *Assessment & Evaluation in Higher Education*, 27(5), 443-454. doi:10.1080/0260293022000009311
- Stevens, D. D., & Levi, A. J. (2013). *Introduction to rubrics: An assessment tool to save grading time, convey effective feedback, and promote student learning*: Stylus Publishing, LLC.