

Fiabilidad y validez de rúbricas para la evaluación de las competencias «Comunicación» y «Diseño y proyecto»

Reliability and validity of scoring rubrics in assessment of competences 'Communication' and 'Design and project'

Fernando Candela^a, Andrea Conchado^a, Lucía Egea^a, Julia Ferrándiz^a, Iván García^a, Lorenzo García^a, Vicent Gassó^a, Victoria Ibáñez^a, Aitziber Irisarri^a, Jaime Nácher-Mestre^a, Armando Nieto^a, Luiz Sanz^a, Jesús Tejel^a

^aCentro Universitario EDEM

Abstract

This paper addresses the analysis of reliability and validity of two rubrics developed by the Universitat Politècnica de València. This analysis is based on the assessment of the presentations made by first year students in EDEM University Center. With this aim, internal consistency and reliability measures have been computed and construct validity has been examined for both rubrics through exploratory and confirmatory factor analysis. Results showed that the rubric for the evaluation of 'Design and project' can be applied to the assessment of these projects, whereas the rubrics for the evaluation of 'Communication' may not be adequate for this context. From a practical point of view, these results have relevant implications for higher education professors involved in the assessment of project - based learning. Meanwhile, these findings have revealed that the analysis of reliability and validity of assessment instruments is essential, when they are applied in different contexts.

Keywords: rubrics, competences, reliability, validity, project – based learning

Resumen

Este trabajo plantea el análisis de fiabilidad y validez de dos rúbricas elaboradas por la Universidad Politècnica de Valencia. Este análisis se basa en la evaluación de las presentaciones del proyecto transversal desarrollado por los alumnos de primer curso en el Centro Universitario EDEM. Para ello se han obtenido medidas de fiabilidad y consistencia interna y se ha analizado la validez de constructo de ambas rúbricas mediante análisis factorial exploratorio y confirmatorio. Los resultados mostraron que la rúbrica para la competencia «Diseño y proyecto» puede aplicarse de forma válida y consistente para la evaluación de estos proyectos, mientras que la rúbrica para la evaluación de la competencia «Comunicación» puede no ser aplicable en este contexto. A nivel práctico, estos resultados tienen importantes implicaciones para los profesores involucrados en la evaluación del aprendizaje orientado a proyectos. Al mismo tiempo, alertan de la importancia de analizar la fiabilidad y validez de los instrumentos de

evaluación, cuando se aplican en contextos distintos a los que fueron creados.

Palabras clave: *rúbricas, competencias, fiabilidad, validez, aprendizaje orientado a proyectos*

1. Introducción

Los últimos dos años han presenciado un creciente interés por las competencias transversales gracias al apoyo institucional de la universidad a través de cursos de formación para el profesorado y el desarrollo de instrumentos de evaluación de competencias. Como medio para formar al alumnado en estas competencias transversales con un enfoque de aprendizaje profundo, la metodología de aprendizaje orientado a proyectos ha mostrado tener gran utilidad y aplicación en este ámbito. Este método de enseñanza fortalece la autoconfianza de los estudiantes, al fomentar que tomen sus propias decisiones y actuaciones de forma independiente (Blumenfeld et al., 1991). Sin embargo, requiere que la aplicación de los conocimientos, habilidades y actitudes guarde semejanzas con las situaciones profesionales reales que tratan de representar, para que realmente el aprendizaje sea más motivador (Brown, Collins y Duguid, 1989). En cualquier caso, esta motivación por el aprendizaje no tendrá sentido si no se identifica el instrumento de medida adecuado para cuantificar la eficacia del aprendizaje orientado a proyectos. Con este objetivo, el Instituto de Ciencias de Educación (ICE) de la Universidad Politécnica de Valencia ha desarrollado rúbricas para la evaluación de trece competencias transversales, que ya se incluyen en las guías docentes de los títulos. En este trabajo se han aplicado dos de estas rúbricas para la evaluación de las competencias adquiridas por los estudiantes de primer curso del Grado en Ingeniería y Gestión Empresarial, en el Centro Universitario EDEM. Los alumnos de esta titulación tienen la responsabilidad de desarrollar un proyecto transversal, con la ayuda y orientación de todos los profesores que imparten docencia en la titulación. Se trata de una iniciativa pionera en el uso de esta metodología, cuya aplicación a toda una titulación de forma sistemática e integrada es difícilmente comparable.

1.1. El Grado en Ingeniería y Gestión Empresarial en el Centro Universitario EDEM

El Centro Universitario EDEM, adscrito a la Universitat de València y a la Universitat Politècnica de València, es una institución académica promovida por la “Escuela de Empresarios, Fundación de la Comunitat Valenciana” (EDEM) en la que se imparten estudios universitarios de carácter oficial de Grado y títulos propios de Postgrado.

EDEM es una fundación sin ánimo de lucro, de carácter privado. Su misión es la formación de empresarios, directivos y emprendedores y el fomento del liderazgo, el espíritu emprendedor y la cultura del esfuerzo. Fue creada hace más de 13 años en el entorno de la

Asociación Valenciana de Empresarios (AVE). Los patronos y miembros de la escuela son empresas de reconocido prestigio.

Por tanto, en aras de cumplir dicha misión, el Centro Universitario EDEM nace con objeto de formar graduados y postgraduados con espíritu emprendedor e intraemprendedor que sean capaces de adaptarse a un mercado laboral en constante cambio y puedan desarrollar proyectos empresariales sostenibles o incorporarse a empresas ya existentes, contribuyendo al desarrollo de nuestra sociedad.

Nuestro modelo educativo se basa en el enfoque práctico y cercano al mundo empresarial, tratando de inculcar los valores, habilidades y conocimientos necesarios en la realidad laboral.

En este contexto surge el Grado en Ingeniería y Gestión Empresarial, titulación oficial adscrita a la Universitat Politècnica de València, con la voluntad de formar a los ingenieros del siglo XXI en conocimientos y habilidades fundamentales para la empresa, para que ocupen posiciones de gestión empresarial con orientación técnica en proyectos globales o como intraemprendedores, desarrollando nuevas áreas de negocio en una empresa.

La titulación, de nueva implantación en el espacio de educación superior español, está verificada por la rama de ingeniería y arquitectura y el plan de estudios se divide en 240 ECTS distribuidos de la siguiente manera:

- 24%: Asignaturas de Ciencias Básicas
- 30%: Asignaturas Tecnológicas
- 46%: Asignaturas de Empresa y Gestión

El alumno de esta titulación debe adquirir las competencias que le permitan ser un interlocutor válido entre todas las áreas funcionales de la empresa (dirección, producción, capital humano, compra, ventas...) para que la tecnología y la organización se relacionen de forma eficaz con el fin de mejorar la competitividad empresarial.

Se pueden identificar como objetivos específicos a conseguir en el Graduado de “Ingeniería y Gestión Empresarial” el desarrollo de las siguientes capacidades:

- Gestionar la innovación y el desarrollo de nuevos productos y servicios teniendo en cuenta la satisfacción de las necesidades del cliente y entendiendo su repercusión en las ventas.
- Gestionar los sistemas de dirección de personas, dentro de un modelo de liderazgo y motivación.
- Poseer los conocimientos tecnológicos necesarios, para mejorar la eficiencia y competitividad de las organizaciones del S.XXI.
- Gestionar la estructura económico-financiera de la empresa.
- Gestionar la apertura de nuevos mercados internacionales.

- Dirigir la empresa desde el punto de vista estratégico, teniendo en cuenta todos sus componentes (cliente, trabajador, proveedor, sociedad, capital).

Asimismo, con el fin de garantizar que los estudiantes poseen una formación integral para el emprendimiento y la dirección de empresas, EDEM ofrece un diploma adicional compuesto por diferentes talleres que se imparten a lo largo de los cuatro años del Grado. Este Diploma de Habilidades Directivas, durante el primer año incluye talleres de inteligencia emocional, gestión del tiempo y trabajo en equipo, así como otras competencias de gran importancia para la formación de los estudiantes.

1.2. Análisis de fiabilidad y validez aplicado a rúbricas

El uso de rúbricas fiables para la evaluación del rendimiento académico aporta interesantes ventajas al profesorado, frente al uso de otros instrumentos de evaluación tradicionales, como los exámenes tipo test o las listas de control (también llamadas listas de verificación). Particularmente las rúbricas tienen un gran potencial para la mejora de la docencia si están bien diseñadas, están centradas en una competencia bien definida y se proporciona a los evaluadores la información necesaria para emplearla de forma consistente (Jonsson y Svingby, 2007). De hecho, algunos estudios apuntan a la necesidad de formar adecuadamente a los evaluadores en el uso de las rúbricas para evitar inconsistencias que disminuyan la fiabilidad y validez del cuestionario.

La fiabilidad es un prerrequisito para la determinación de la validez de una rúbrica. A pesar de que una medición válida es necesariamente fiable, la afirmación contraria no siempre es verdadera. En otras palabras, una medición fiable no tiene porqué ser válida. En este caso, puede ocurrir que efectivamente la rúbrica muestre puntuaciones similares en distintas condiciones (distintos evaluadores o momentos de evaluación). Sin embargo, lo que miden estas puntuaciones puede estar muy alejado de lo que en realidad se pretende medir (Moskal y Leydens, 2000). Para ello, el análisis de la validez constituye un elemento esencial a la hora de decidir si una rúbrica es adecuada para medir determinadas competencias (Rezaei y Lovorn, 2010). No obstante, es necesario ser consciente de las valoraciones sobre la fiabilidad de una rúbrica se obtienen siempre a partir de un determinado conjunto de puntuaciones. Es decir, las afirmaciones sobre la fiabilidad de una rúbrica en un determinado contexto pueden no ser aplicables en otras circunstancias (Morales, 2006).

Un estudio publicado en 2013 muestra cómo distintos factores relacionados con el empleo de la rúbrica pueden afectar al efecto que tienen en la mejora del rendimiento y la implicación del alumno en el aprendizaje (Panadero y Jonsson, 2013). Este trabajo enfatiza en el potencial de las rúbricas para influir positivamente en el aprendizaje de los estudiantes, a través de la evaluación formativa, en detrimento de la evaluación sumativa que empleada tradicionalmente en docencia. Según este estudio, determinados factores puede moderar el efecto de las rúbricas en el aprendizaje, como las habilidades

metacognitivas o el género del estudiante, por lo que recomienda analizar la influencia de las características personales en la evaluación formativa. De hecho, otro estudio posterior profundiza en las diferencias observadas en las evaluaciones e interpretaciones hechas por estudiantes y profesores, ante los mismos ensayos y redacciones elaboradas por estudiantes matriculados en un curso de escritura (Li y Lindsey, 2015).

Al analizar la fiabilidad de una rúbrica, debemos estudiar la consistencia interna del instrumento, y más concretamente mediciones de la homogeneidad del constructo que se mide. Este enfoque surge de la teoría clásica de la fiabilidad y se basan fundamentalmente en el coeficiente α de Cronbach y las correlaciones inter – ítem. Ambas medidas de fiabilidad aportan información sobre la homogeneidad del constructo, en este caso sobre la competencia, que se pretende medir. Este concepto se presta a equívocos en muchas ocasiones, pues tienen mucha relación con la unidimensionalidad de la rúbrica. La unidimensionalidad se refiere a la estructura interna del instrumento (Morales, 2006), e indica semejanza conceptual entre los ítems de la rúbrica, o existencia de un único factor muy claro.

Para analizar en qué medida los ítems de una rúbrica están midiendo lo mismo, es decir, tienen algo en común, se utiliza habitualmente el coeficiente α de Cronbach (1951). Este coeficiente representa la proporción entre varianza verdadera y varianza total. En esta definición el término varianza verdadera se debe a que los estudiantes difieren en aquello que tienen en común con los ítems de la rúbrica. Se ha propuesto un valor mínimo de 0.70 de fiabilidad para este coeficiente (Nunnally y Bernstein, 1994) pero este valor depende en cualquier caso del uso que se hará del instrumento de medida. Este coeficiente tiene la ventaja de su claridad conceptual y es un método muy establecido para estimar la fiabilidad. Sin embargo, posee dos limitaciones de gran calado como su dependencia del número de ítems y su pérdida de interpretabilidad cuando se aplica a constructos multidimensionales.

Por ello se recomiendan otros métodos complementarios para la evaluación de la fiabilidad de las rúbricas, que permitan verificar la unidimensionalidad de la competencia que se pretende medir, como las correlaciones ítem – total basadas en el análisis de ítems. Estas correlaciones están calculadas en base a la relación de cada ítem con la puntuación total obtenida con la rúbrica completa, excluyendo el ítem en cuestión. En estas medidas se ha sugerido una puntuación mínima de 0.20 para considerar que existe consistencia interna entre los ítems.

Un enfoque alternativo para comprobar la unidimensionalidad de una escala es utilizar análisis factorial exploratorio. Este procedimiento se basa en el análisis de la varianza común a los ítems de la rúbrica, y utiliza desde el principio los pesos de los ítems en cada factor. Así, esta técnica de análisis permite estudiar la correlación ítem – factor, reemplazando a la correlación ítem – total. Al emplear análisis factorial exploratorio para

validar la unidimensionalidad de una rúbrica, cabe esperar que no todos los ítems estén igualmente relacionados con la competencia que se mide. En este caso, se centra la atención en el primer factor pues extrae la máxima cantidad de varianza común a todos los ítems, y en este sentido, es el más informativo (Peña, 2014). Al estudiar cómo los ítems cargan sobre este primer factor, se ha establecido un peso mínimo en torno a 0.30 (Hair, Black, Babin, Anderson y Tatham, 2006), teniendo en cuenta también el mismo contenido de los ítems. Por otro lado, independientemente de que los resultados del análisis sean interpretables, la rotación de los factores siempre aporta información de interés para valorar la competencia que se mide a través de la rúbrica. Entre los métodos disponibles para realizar la rotación, el procedimiento de rotación ortogonal Varimax asegura que los pesos de las variables en los factores tengan la mayor varianza posible. Esto significa que los ítems de la rúbrica tienden a tener correlaciones altas con unos factores y muy bajas en otros. Por tanto, ésta es la solución más útil para diferenciar conjuntos de ítems, y verificar así la unidimensionalidad de la rúbrica.

No obstante, la aplicación de estas técnicas se limita a la evaluación de la fiabilidad de la rúbrica. Para estudiar en qué medida una rúbrica realmente mide lo que pretende medir es necesario analizar la validez del instrumento. Este análisis puede realizarse desde distintos enfoques, abordando distintos tipos de validez: validez de contenido, validez de constructo y validez de criterio.

La validez de contenido indica el grado en que los ítems seleccionados en la rúbrica constituyen una definición adecuada de la competencia. Es decir, refleja si los ítems seleccionados cubren todos los aspectos relativos a la definición del alcance de la competencia. En este sentido, la validez de contenido puede relacionarse con los resultados del aprendizaje vinculados a la competencia, y se basa en gran medida en el juicio experto de los profesores implicados en la impartición de la asignatura.

La validez de constructo analiza en qué medida subyace un factor latente a los ítems seleccionados en la rúbrica. Un constructo es un concepto teórico, que en el caso de las rúbricas hace referencia a las competencias medidas. Al tratarse de un factor latente, la única forma de medir las competencias es a través de los ítems definidos en la rúbrica, que se definen como variables manifiestas. El análisis factorial confirmatorio, como parte de los modelos de ecuaciones estructurales, permite confirmar si los distintos ítems de una rúbrica están midiendo una misma competencia. O si, por el contrario, reflejan el efecto de varios componentes. La clave de la aplicación de este análisis a la validación de rúbricas reside en la especificación de un modelo de medida unidimensional. La validación de este modelo se basa en múltiples criterios. El primero de ellos es el valor del estadístico χ^2 y su nivel de significación asociado. Sin embargo, puesto que este estadístico es sensible al tamaño de la muestra (Cheung y Renswold, 2002), se recomienda emplear índices de bondad de ajuste alternativos (Hu y Bentler, 1999), como los índices NFI y NNFI (*Normed and Non – Normed Fit Indexes*), el índice CFI (*Comparative Fit Index*) y el índice RMSEA (*Root –*

Mean – Square – Error of Approximation) (Bentler, 1990; Steiger, 1990). El índice CFI varía entre 0 y 1, de modo que los valores superiores a 0.90 y 0.95 indican un ajuste aceptable del modelo, o excelente, respectivamente. Por el contrario, son preferibles valores pequeños para el índice RMSEA. Valores inferiores a 0.05 indican un ajuste excelente del modelo, mientras que valores inferiores a 0.08 indican ajuste aceptable (Browne y Cudeck, 1993).

2. Objetivos

El objetivo general de este trabajo es analizar la fiabilidad y validez de dos rúbricas elaboradas por el Instituto de Ciencias de Educación de la Universidad Politécnica de Valencia a través de la evaluación de las presentaciones del proyecto transversal desarrollado por los alumnos de primer curso en el Centro Universitario EDEM. Este objetivo general se concreta en los siguientes objetivos específicos:

- Estudiar la variabilidad debida a los distintos evaluadores que hicieron uso de las rúbricas
- Estimar la consistencia interna y fiabilidad de los ítems que componen cada rúbrica
- Valorar la validez de las rúbricas en su aplicación a la evaluación de competencias en los proyectos desarrollados por los estudiantes del Centro Universitario EDEM.

3. Desarrollo de la innovación

3.1. Descripción de la innovación

La innovación docente objeto de estudio ha consistido en plantear a los alumnos de primer curso del «Grado en Ingeniería y Gestión Empresarial» un proyecto anual, transversal a todas las asignaturas, en el que aplicar parte de lo aprendido en clase, con el objetivo final de realizar una competición entre equipos. El proyecto anual se ha implantado en el primer curso (2015-2016) de la primera promoción del Grado en Ingeniería y Gestión Empresarial.

Los objetivos de implantar un proyecto anual que fuera transversal a todas las asignaturas han sido varios:

- Motivar a los alumnos:
 1. Con una competición por equipos
 2. Viendo la aplicación práctica e inmediata de lo aprendido en clase
 3. Despertando la curiosidad por aprender
- Aprendizaje de competencias y habilidades necesarias en su futura experiencia profesional:
 1. Trabajo en equipo
 2. Resolución de problemas de manera autónoma
 3. Presentaciones eficaces. Comunicación y síntesis de resultados.
- Comprender la importancia de la visión global de los problemas.

Fase I: Diseño

Académico

El proyecto se ha relacionado con cada una de las asignaturas del Plan de Estudios y se incluye dentro del sistema de evaluación de la Guía Docente con un peso del 10% de la nota final, correspondiente a la evaluación continua.

La definición del enunciado del proyecto se ha realizado por parte del claustro docente a través de varias reuniones de coordinación, hasta encontrar un tema en el que se pudieran recoger parte de los contenidos de todas las asignaturas que en este caso eran las siguientes:

- 1º Semestre: Física I, Cálculo, Química, Empresa e Informática.
- 2º Semestre: Física II, Ecuaciones diferenciales, Álgebra, Biología, Expresión gráfica y Economía.

Cuando se llegó a un consenso, cada profesor definió la entrega que los alumnos debían realizar en su asignatura y la plasmaba en la Guía Docente. El objetivo de obtener una nota académica ha ayudado a mantener la implicación del alumnado en el mismo, ya que al finalizar el curso se evaluaba un total de 6 ECTS.

Se diseñó un documento semestral donde se plasmaba el calendario de entrega de todas las asignaturas, así como las especificaciones requeridas. Si bien es cierto que el proyecto tiene un alcance anual, se ha definido una presentación a la finalización del primer semestre (Diciembre de 2015), con el objetivo de recapitular el estado de desarrollo de los proyectos, momento en que se recopilaron las evaluaciones del presente trabajo.

Gestión alumnado

Los alumnos han trabajado en ocho grupos de cuatro personas. Los grupos se configuraron por el Centro Universitario EDEM, con el criterio de heterogeneidad dentro del grupo y homogeneidad entre grupos en función del perfil de los alumnos.

Fase II: Implementación

Académica

Desde el punto de vista académico, las fases de implementación han sido las siguientes:

En primer lugar, por parte de EDEM se realizó una explicación general a todos los alumnos con el objetivo de explicar la motivación de trabajar en un proyecto anual, centrar expectativas y definir en términos generales el proyecto en cuestión y la operativa de trabajo (calendarización y especificaciones).

Seguidamente, cada profesor, al inicio de la asignatura, explicó la parte del proyecto relativa a su asignatura, así como los criterios de evaluación y fechas de entrega. A lo largo del semestre, se realizaron varias reuniones de coordinación entre los profesores para evaluar la evolución del proyecto y ajustar las tareas encomendadas.

Los alumnos mostraron dudas en la realización de los hitos evaluables definidos en cada asignatura. Una de las razones fue el calendario de impartición de los contenidos necesarios para la realización de las actividades y la falta de autonomía en la resolución de problemas con varias soluciones posibles. Por ello, se decidió desglosar las entregas en pequeños sub-hitos para acompañar a los alumnos en el proceso de aprendizaje.

Gestión alumnado

Los alumnos reaccionaron favorablemente a realizar este tipo de actividad dentro de la asignatura, pero surgieron algunos inconvenientes en la implementación, debido a que supone una carga de trabajo adicional al estudio y a la dificultad de coordinación con los compañeros del grupo. Este inconveniente origina entregas en el último momento y conflictos entre alumnos por la diferente implicación de algunos de ellos. Con los grupos de alumnos que expusieron sus conflictos a EDEM, se realizaron reuniones individuales y grupales para mejorar las competencias de trabajo en equipo, así como sesiones de formación relacionadas con estos aspectos.

Fase III: Evaluación de resultados

Académica

La evaluación se realizó por parte de cada profesor a nivel grupal, es decir, por cada entrega de los alumnos, se otorgaba una calificación por grupo, que influía directamente en la nota individual de cada alumno en la asignatura (10% de la nota final) correspondiente a la evaluación continua. Adicionalmente, para evaluar la implicación de cada miembro de grupo se pidió a los estudiantes que valoraran a sus compañeros de equipo, generando un factor de ponderación que se aplica sobre la nota final.

Gestión alumnado

Además, con el fin de generar el espíritu competitivo y fomentar habilidades de síntesis y comunicación los alumnos debían presentar los resultados obtenidos en cada asignatura durante el semestre en presencia de un jurado compuesto por profesores de EDEM. Para realizar esta evaluación se emplearon las rúbricas desarrolladas por el Instituto de Ciencias de la Educación de la UPV que se describen a continuación.

3.2. Instrumentos: Rúbricas del ICE

Para la medición de la adquisición de competencias adquiridas en la presentación de resultados del primer semestre, se emplearon dos rúbricas desarrolladas por el Instituto de Ciencias de la Educación de la UPV. En el caso concreto de la innovación que nos ocupa, se decidió emplear las rúbricas correspondientes a las competencias transversales:

- CT-05. Diseño y proyecto. Nivel de dominio I
- CT-08. Comunicación efectiva. Nivel de dominio I

3.3. Análisis de datos

El análisis de datos que se muestra en este trabajo se centra en las evaluaciones que cuatro profesores hicieron a partir de la presentación de resultados tras la finalización del primer semestre. Esta evaluación tuvo lugar en diciembre de 2015. Las evaluaciones se hicieron por grupos y no a nivel individual.

En primer lugar se muestran los resultados descriptivos de los ítems correspondientes a cada rúbrica, incluyendo indicadores de asimetría y curtosis para la verificación de los supuestos de normalidad univariante en los datos. Estos descriptivos se complementan con una representación gráfica de las puntuaciones asignadas por cada evaluador, con el fin de analizar la variabilidad inter – jueces. El hecho de que distintos evaluadores puntúen de forma similar las mismas presentaciones de proyectos, indica que el instrumento es fiable respecto a esta fuente de incertidumbre.

A continuación, se exponen los resultados del análisis de ítems en base a las correlaciones ítem – total para cada rúbrica. Igualmente se muestran los valores del coeficiente α de Cronbach. Estos indicadores aportan información sobre la fiabilidad de cada rúbrica, aplicada a las puntuaciones obtenidas en las presentaciones de los proyectos elaborados por los estudiantes del primer curso de Ingeniería y Gestión Empresarial en EDEM.

Seguidamente se ha analizado la validez de constructo de cada rúbrica utilizando análisis factorial exploratorio y confirmatorio. Previamente a la extracción de factores mediante componentes principales, se ha verificado el valor de la medida KMO. La matriz de cargas factoriales obtenida se ha transformado utilizando rotación varimax. El modelo de medida planteado en la especificación para el análisis factorial confirmatorio ha planteado la existencia de un único factor subyacente a los ítems de cada rúbrica. En otras palabras, se ha especificado que cada rúbrica mide solo un constructo: las competencias de comunicación y diseño y proyecto, correspondientemente. Para cada rúbrica se muestran los índices de bondad de ajuste y las ecuaciones estandarizadas que indican el peso de cada ítem en cada factor latente.

4. Resultados

En términos generales se observaron puntuaciones superiores en la competencia «Comunicación», respecto a los promedios obtenidos en la competencia «Diseño y proyecto». Como puntos fuertes de los estudiantes se detectó una disposición favorable hacia la comunicación, capacidad para la transmisión de información relevante y uso del lenguaje con corrección gramatical. Por el contrario, la asignación de plazos para el desarrollo del proyecto y la identificación de riesgos inherentes al proyecto parecen ser los puntos débiles de los estudiantes. Estos resultados pueden deberse a que la formación de los

estudiantes para el desarrollo de proyectos se encontraba todavía en proceso, en el momento en que presentaron sus primeros avances en el proyecto.

Tabla 1. Resultados descriptivos de los ítems contenidos en ambas rúbricas

		Media	Desv. típ.	Asimetría	Curtosis
Comunicación	C1 - Muestra una disposición favorable hacia la comunicación	3,27	0,46	1,10	-0,89
	C2 - Transmite información relevante y sabe responder a las preguntas que se formulan	3,12	0,33	2,61	5,44
	C3 - Utiliza el lenguaje con corrección gramatical:semántica, sintáctica y ortográfica	3,05	0,59	0,00	0,35
	C4 - Las exposiciones están estructuradas de modo coherente	3,00	0,55	0,00	0,99
	C5 - Utiliza medios de apoyo pertinente	3,00	0,65	0,00	-0,28
	C6 - Su comunicación no verbal transmite tranquilidad	2,81	0,81	-0,85	0,93
Diseño y proyecto	DP1 - Justifica razonadamente la necesidad del proyecto	2,86	0,85	0,29	-1,59
	DP2 - Establece unos objetivos claros del proyecto	2,75	0,85	-0,61	0,24
	DP3 - Propone las acciones para la consecución de los objetivos (y asigna responsables en el caso de trabajar en grupo)	2,47	0,77	-0,31	-0,10
	DP4 - Desarrolla las acciones para la consecución de objetivos	2,33	0,91	0,30	-0,40
	DP5 - Identifica posibles riesgos inherentes al proyecto	2,26	0,99	-0,21	-1,40
	DP6 - Asigna plazos necesarios para completar las acciones previstas (y asigna responsables en el caso de trabajar en grupo)	1,88	0,93	0,26	-1,91

Apenas se encontraron diferencias significativas en las valoraciones realizadas por distintos profesores, o jueces, respecto a cada uno de los ítems incluidos en la rúbrica sobre comunicación. Este resultado indica que la aplicación de esta rúbrica para la evaluación de la competencia comunicación resulta fiable para los estudiantes de este centro. El único ítem en el que se detectaron discrepancias fue el C6 «Su comunicación no verbal transmite tranquilidad». Cabe la posibilidad de que el término «tranquilidad» deba ser redactado con mayor precisión para evitar distintas interpretaciones.

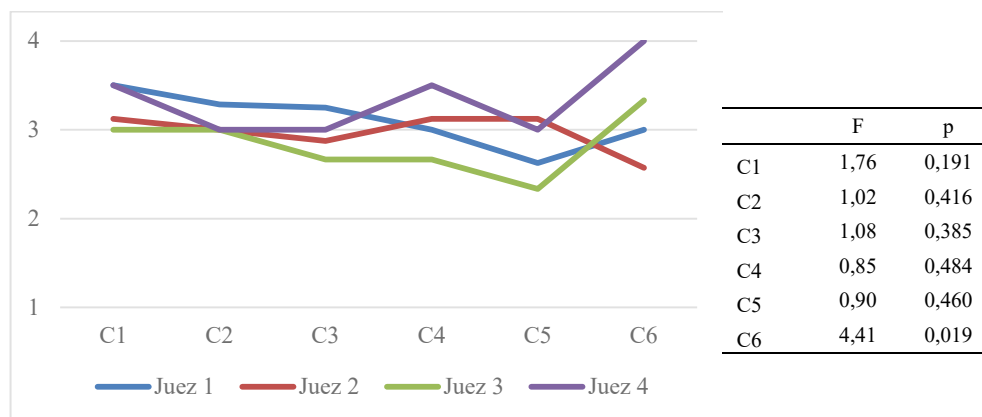


Fig. 1 Variabilidad inter – jueces para la competencia «Comunicación». Tabla ANOVA para el análisis de

variabilidad en las puntuaciones, debida al factor juez.

Por el contrario, los ítems de la rúbrica «Diseño y proyecto» presentaron marcadas diferencias, que en todos los casos menos uno resultaron estadísticamente significativas. En consecuencia, puede decirse que existen fuertes discrepancias en la interpretación de los ítems y la forma de valorarlos cuando se aplica esta rúbrica a la evaluación de los proyectos desarrollados por los alumnos del centro.

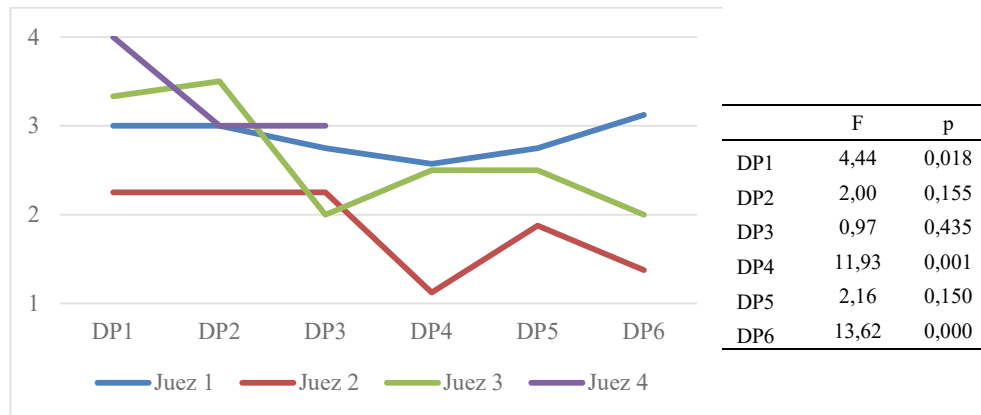


Fig. 2 Variabilidad inter –jueces para la competencia «Diseño y proyecto». Tabla ANOVA para el análisis de variabilidad en las puntuaciones, debida al factor juez.

Los resultados del análisis de ítems aplicado a la rúbrica para la competencia «Comunicación» mostraron la baja consistencia interna del instrumento ($\alpha = 0.545$). Las correlaciones ítem – total para los ítems C2 y C6 presentaron valores cercanos a cero, indicando así la debilidad de la relación entre dichos ítems y la competencia que pretende medir esta rúbrica. En otras palabras, el hecho de transmitir información relevante y utilizar medios de apoyo pertinentes parece ser independiente de poseer la competencia de comunicación. Ambos ítems pueden relacionarse con otras habilidades distintas a la comunicación, como el dominio técnico del área de estudio o la capacidad tecnológica de preparar elaboradas presentaciones con medios digitales.

Tabla 2. Análisis de ítems para la competencia «Comunicación»

	Correlación elemento-total corregida	Correlación múltiple al cuadrado	Alfa de Cronbach si se elimina el elemento
C1 - Muestra una disposición favorable hacia la comunicación	,463	,535	,429
C2 - Transmite información relevante y sabe responder a las preguntas que se formulan	-,081	,558	,605
C3 - Las exposiciones están estructuradas de modo coherente	,463	,388	,417
C4 - Utiliza el lenguaje con corrección gramatical: semántica, sintáctica y ortográfica	,406	,633	,440

C5 - Su comunicación no verbal transmite tranquilidad	,343	,278	,474
C6 - Utiliza medios de apoyo pertinente	,141	,123	,579

Por el contrario, la rúbrica para la evaluación de la competencia «Diseño y proyecto» obtuvo un elevado valor del coeficiente α de Cronbach ($\alpha = 0.862$). Ningún ítem mostró correlaciones ítem – total inferiores a 0.4 lo cual indica que todos los ítems contribuyen de forma significativa a medir la misma competencia. En otras palabras, la aplicación de esta rúbrica para la evaluación de la competencia «Diseño y proyecto» es fiable para las puntuaciones obtenidas por los estudiantes del centro.

Tabla 3. Análisis de ítems para la competencia «Diseño y proyecto»

	Correlación elemento-total corregida	Correlación múltiple al cuadrado	Alfa de Cronbach si se elimina el elemento
Justifica razonadamente la necesidad del proyecto	,619	,664	,845
Establece unos objetivos claros del proyecto	,756	,785	,820
Propone las acciones para la consecución de los objetivos (y asigna responsables en el caso de trabajar en grupo)	,642	,743	,842
Asigna plazos necesarios para completar las acciones previstas (y asigna responsables en el caso de trabajar en grupo)	,755	,748	,820
Desarrolla las acciones para la consecución de objetivos	,738	,802	,824
Identifica posibles riesgos inherentes al proyecto	,459	,458	,878

A continuación se aplicaron técnicas de análisis factorial para analizar la validez de constructo de ambas rúbricas. La medida KMO para la rúbrica Comunicación presentó un valor muy reducido que indicaba baja correlación entre los ítems de dicha rúbrica, como ya había anticipado el análisis de ítems ($KMO = 0.448$). Este resultado fue confirmado mediante la prueba de esfericidad de Barlett, que resultó no significativa ($\chi^2 = 22.4, p = 0.099$).

A pesar de estos resultados, se aplicó análisis factorial exploratorio y se extrajeron tres componentes principales, que explicaban el 77.4% de la variabilidad de los datos, según el criterio de Kaiser que retiene todos los autovalores superiores a la unidad (Factor 1: 2.095; Factor 2: 1.549; Factor 3: 1.001). La primera componente tan solo explicó el 34.9%

Tabla 4. Matriz de cargas factoriales rotadas para la competencia «Comunicación»

	Factor		
	1	2	3
Muestra una disposición favorable hacia la comunicación	,394	,831	,070
Transmite información relevante y sabe responder a las preguntas que se formulan	-,337	,874	-,022
Las exposiciones están estructuradas de modo coherente	,719	,274	,070
Utiliza el lenguaje con corrección gramatical: semántica, sintáctica y ortográfica	,919	-,112	-,125
Su comunicación no verbal transmite tranquilidad	,644	-,097	,224
Utiliza medios de apoyo pertinente	,071	,034	,983

Para la especificación del modelo de medida en análisis factorial confirmatorio, el test de Mardia indicó que los ítems cumplían el supuesto de normalidad multivariante requerido para la ejecución de análisis factorial confirmatorio (Estadístico de Mardia = $1.54 \leq 5$). El modelo unidimensional se ajustó de forma excelente a los datos (NFI = 0.861, NNFI = 1.652, CFI = 1.000, $\chi^2 = 3.78$, $p = 0.876$, RMSEA = 0.000 [0.000, 0.142]). Sin embargo, la varianza del factor latente arrojó un valor muy próximo a cero, lo que constituye una estimación infractora (Varianza del factor: 0.019). Asimismo, los coeficientes de regresión entre cada ítem y la competencia representada en el factor latente no resultaron significativos en ningún caso. Ambos resultados no son suficientes para invalidar los resultados del modelo, pero sí para señalar que deben ser interpretados con cautela.

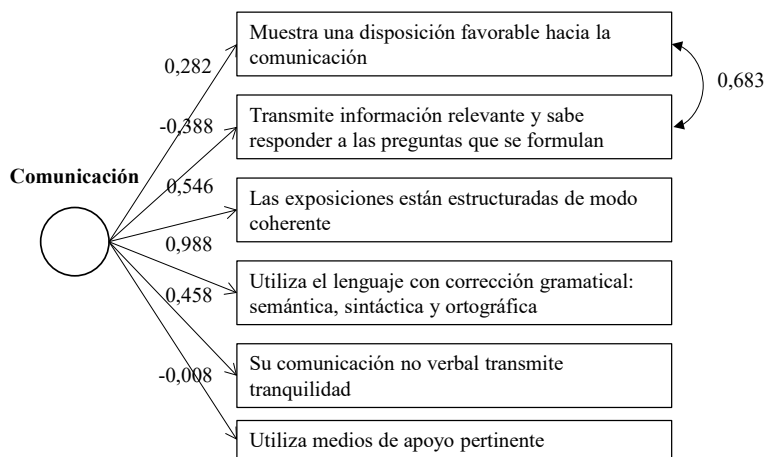


Fig. 3 Estimación de los parámetros del modelo de medida para la competencia «Comunicación»

El valor de la medida KMO para la competencia «Diseño y proyecto» fue superior al obtenido anteriormente ($KMO = 0.666$), pero la prueba de esfericidad de Barlett sí resultó significativa en este caso ($\chi^2 = 55.8$, $p = 0.000$). No obstante, en este caso el análisis factorial exploratorio sí extrajo una única componente principal, confirmando la unidimensionalidad del constructo. Es decir, sí puede afirmarse para esta competencia que existe un único factor subyacente a todos los ítems que componen la rúbrica para «Diseño y proyecto». Esta componente única explicó el 60.5% de la variabilidad de los datos según el criterio de Kaiser que retiene todos los autovalores superiores a la unidad (Factor 1: 3.635)

Tabla 5. Matriz de cargas factoriales rotadas para la competencia «Diseño y proyecto»

	Factor
	1
Justifica razonadamente la necesidad del proyecto	,741
Establece unos objetivos claros del proyecto	,869
Propone las acciones para la consecución de los objetivos (y asigna responsables en el caso de trabajar en grupo)	,778
Asigna plazos necesarios para completar las acciones previstas (y	,827

asigna responsables en el caso de trabajar en grupo)	
Desarrolla las acciones para la consecución de objetivos	,843
Identifica posibles riesgos inherentes al proyecto	,575

El test de Mardia indicó en este caso que los ítems no cumplían el supuesto de normalidad multivariante requerido para la ejecución de análisis factorial confirmatorio (Estadístico de Mardia = $7.49 \geq 5$) por lo que fue necesario recurrir al método de estimación robusto. El modelo unidimensional se ajustó de forma excelente a los datos (NFI = 0.883, NNFI = 0.928, CFI = 0.967, $S - B\chi^2 = 9.09$, $p = 0.245$, RMSEA = 0.137 [0.000, 0.345]) sin estimaciones infractoras que invalidaran la interpretación de resultados (Varianza del factor: 0.239). Todos los coeficientes de regresión del modelo de medida resultaron significativos.

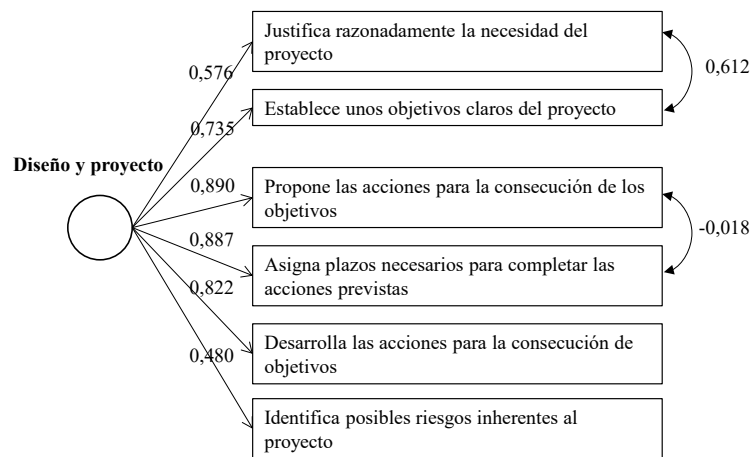


Fig. 4 Estimación de los parámetros del modelo de medida para la competencia «Diseño y proyecto»

5. Conclusiones

El presente estudio ha mostrado que la rúbrica para la competencia «Diseño y proyecto» puede aplicarse de forma válida y fiable para la evaluación de los proyectos elaborados por los estudiantes en el Grado de Ingeniería y Gestión Empresarial del Centro Universitario EDEM. Por el contrario, la aplicación de la rúbrica para la competencia «Comunicación» no parece ser procedente para los alumnos de este centro. Esta discrepancia puede deberse al particular enfoque del centro hacia la formación empresarial de los estudiantes, centrado en el desarrollo de determinadas habilidades de gestión. A nivel práctico, este trabajo tiene importantes implicaciones para el profesorado involucrado en la evaluación del aprendizaje orientado a proyectos. A nivel académico, se evidencia la necesidad de revisar cualquier instrumento de evaluación si se pretende aplicar en contextos distintos a los que fueron diseñados. Tal y como se ha descrito anteriormente, la fiabilidad de una rúbrica no es inherente a sí misma, sino a las puntuaciones en base a las cuales se ha validado. En aquellos casos en que se considere necesario, los profesores deberán plantearse la modificación de las rúbricas existentes e incluso la elaboración de rúbricas propias,

adecuadas al alcance y nivel de rendimiento que se espera obtener de los alumnos al finalizar una actividad de enseñanza – aprendizaje.

6. Referencias

- BENTLER, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238.
- BLUMENFELD, P., SOLOWAY, E., MARX, R., KRAJCIK, J., GUZDIAL, M., y PALINCSAR, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. *Educational Psychologist*, 26 (3y4), 369-398.
- BROWN, J. S., COLLINS, A. y DUGUID, P. (1989) Situated cognition of learning. *Educational Researcher*, 18, 32-42.
- BROWNE, M W y CUDECK, R. (1993). Alternative ways of assessing model fit. In *Testing structural equation models* (Sage., pp. 136–162). Newbury Park, CA.
- CHEUNG, G. W., y RENSWOLD, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9, 233–255.
- CRONBACH, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297–334.
- HAIR, J. F.; BLACK, W. C.; BABIN, B. J.; ANDERSON, R. E. y TATHAM, R. L. (2006). *Multivariate data analysis*. Upper Saddle River, NJ: Pearson Prentice Hall.
- HU, L., y BENTLER, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6, 1–55.
- JONSSON, A., y SVINGBY, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2(2), 130–144.
- LI, J., y LINDSEY, P. (2015). Understanding variations between student and teacher application of rubrics. *Assessing Writing*, 26, 67–79. doi:10.1016/j.asw.2015.07.003.
- MORALES, P. (2006). *Medición de actitudes en psicología y educación*. Madrid: Universidad Pontificia de Comillas.
- MOSKAL, J., y LEYDENS, B. (2000). Scoring rubric development: validity and reliability. *Practical Assessment, Research and Evaluation*, 7(10), 71-81.
- NUNNALLY, J. C., y BERNSTEIN, I. H. (1994). The theory of measurement error. *Psychometric Theory*, 209–247.
- PANADERO, E., y JONSSON, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, 9, 129–144.
- PEÑA, D. (2014). *Análisis de datos multivariantes*. Madrid: McGraw-Hill / Interamericana de España.

REZAEI, A. R., y LOVORN, M. (2010). Reliability and validity of rubrics for assessment through writing. *Assessing Writing*, 15(1), 18–39.

STEIGER, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, 25, 173–180.