



Consistencia y concordancia entre profesorado y alumnado en el uso de rúbricas para desarrollo y evaluación de trabajos escritos en la materia biomecánica deportiva

Consistency and agreement between professors and students in the employment of rubrics for the development and assessment of written projects in the subject of sport biomechanics

Jose Luis Sanchez-Jimenez^a, Joaquín Martín Marzano-Felisatti^a, Inma Aparicio^a, Alberto Encarnación-Martínez^a, Irene Jiménez-Pérez^a, Pedro Pérez-Soriano^a y Jose Ignacio Priego-Quesada^a

^a Grupo de Investigación en Biomecánica aplicada al Deporte (GIBD), Departamento de Educación Física y Deportiva, Facultad de Ciencias de la Actividad Física y el Deporte, Universitat de València, (josanji2@alumni.uv.es, ORCID), (marzano@alumni.uv.es, ORCID), (inmaculada.aparicio@uv.es, ORCID), (alberto.encarnacion@uv.es, ORCID), (ijimenez.gibd@gmail.com, ORCID), (pedro.perez-soriano@uv.es, ORCID) y (j.ignacio.priego@uv.es, ORCID).

How to cite: Jose Luis Sanchez-Jimenez, Joaquín Martín Marzano-Felisatti e Inma Aparicio. 2022. Consistencia y concordancia entre profesorado y alumnado en el uso de rúbricas para desarrollo y evaluación de trabajos escritos en la materia biomecánica deportiva. En libro de actas: *VIII Congreso de Innovación Educativa y Docencia en Red*. Valencia, 6 - 8 de julio de 2022. <https://doi.org/10.4995/INRED2022.2022.15867>

Abstract

The implication of the students in assessment process has grown in Superior Studies through the use of rubrics. However, the design of the instruments is a complex task and it must be adapted to the subject with the aim to ensure an optimal evaluation. The objective of this study was to identify differences in the results obtained between self-assessment and teacher assessment and to analyse the inter-evaluative reliability. 48 students participated in the study. At the end of practice session in Biomechanical subject the students prepared a task according to the items included in the rubric and it was delivered together with self-assessment rubric. After that, professors assessed the tasks with the same rubric. The mark obtained by student assessment was higher than the provided by the professor assessment. 67% of the items present a good or excellent repeatability but the remaining 33% showed a regular repeatability. It suggests the need to improve the item description accuracy in the different grades of achieve.

Keywords: learning, academic performance, self-assessment

Resumen

La implicación del alumnado en el proceso de evaluación ha aumentado en las Enseñanzas Superiores mediante el uso de rúbricas. Sin embargo, el diseño del instrumento es una tarea compleja y se debe adecuar a cada asignatura para garantizar una adecuada evaluación. El objetivo del estudio ha sido identificar posibles diferencias en los resultados de las

evaluaciones realizadas por el profesorado y el alumnado, así como analizar la fiabilidad inter-evaluador de la rúbrica diseñada. En el estudio participaron un total de 48 estudiantes. Al finalizar la sesión práctica de la materia de Biomecánica, el alumnado realizó una memoria de la práctica y lo entregó junto a la rúbrica proporcionada para su auto evaluación. Posteriormente, el profesorado evaluó dichos trabajos con la misma rúbrica. Las calificaciones realizadas por el alumnado fueron mayores que las del profesorado en todos los ítems. La repetibilidad intraevaluador fue excelente o buena en el 67% de los ítems, mientras que el 33% restante obtuvo una repetibilidad regular. Por lo tanto, se requiere un aumento de la precisión en la descripción de los distintos niveles de logro.

Palabras clave: *aprendizaje, rendimiento académico, autoevaluación.*

1. Introducción

Los métodos de evaluación en el contexto de la Educación Superior han evolucionado hacia modelos en los cuales el alumnado se ve implicado en el proceso de evaluación, lo que ha provocado un aumento de la variedad de modalidades de evaluación (Berrocoso y Gómez, 2014), de acuerdo a lo establecido en el Real Decreto 1791/2010, de 30 de diciembre, por el que se aprueba el Estatuto del Estudiante Universitario, nº 318, donde se determina que se tenderá a emplear la evaluación continua y por lo tanto, la responsabilidad en el proceso evaluador será compartida entre el alumnado y el profesorado. No obstante, en las universidades españolas continúa predominando un discurso tradicional donde se muestra el personal docente como un agente calificador, al alumnado como receptor de las consecuencias de la evaluación y el examen escrito como la principal técnica evaluativa (Ibarra y Rodríguez, 2010). En este sentido, los métodos con estas características son percibidos por el alumnado como pruebas a superar y no como una oportunidad de aprendizaje (Berrocoso y Gómez, 2014). Por lo tanto, debido a que el método de evaluación influye en el proceso de aprendizaje (Rodríguez y Estrada, 2007) y en la percepción del alumnado que participa en él (Berrocoso y Gómez, 2014), se debe tener en cuenta cuando se elabora el proceso de enseñanza orientándolo también a la evaluación de competencias.

Para garantizar que la evaluación sea formativa, el estudiante debe participar en la tarea de evaluación (Gallego y Raposo, 2014) y, de este modo, el uso de rúbricas como herramienta evaluativa se presenta como una buena estrategia para involucrar al alumnado. Una rúbrica es un elemento que proporciona diferentes niveles de logro para un conjunto de criterios o dimensiones en una situación dada (Allen y Tanner, 2006), permitiendo diferenciar los diferentes niveles de calidad y las expectativas para cada calificación (Andrade y Du, 2019). Por este motivo, la descripción adecuada de cada uno de los logros establecidos en las diferentes dimensiones es imprescindible para obtener un uso adecuado (Allen y Tanner, 2006). Las rúbricas son un elemento esencial en la evaluación formativa ya que se consideran como instrumento de interpretación de los resultados de la evaluación permitiendo la evaluación del progreso y orientada al aprendizaje (Gallego y Raposo, 2014). Además, se ha observado como muy enriquecedora la participación del estudiante en el proceso evaluador, y una alta satisfacción en su empleo (Gallego y Raposo, 2014). Uno de los aspectos positivos del uso de rúbricas es que el alumnado conoce cuales son las características que debe cumplir un trabajo para una calificación dada y facilita el desarrollo de tareas con las que todavía no se encuentran familiarizados (Allen y Tanner, 2006).

La aplicación de rúbricas en el proceso de evaluación ha mostrado resultados positivos en diferentes competencias. Respecto al alumnado, conocer los ítems de evaluación de la tarea proporciona una mayor comprensión de la misma y se obtienen mejores calificaciones (Andrade y Du, 2019). En ese sentido, les

permite identificar los elementos críticos, autoevaluar su progreso de forma periódica y reducir la ansiedad (Andrade y Du, 2019; Berrocoso y Gómez, 2014; Panadero y Jonsson, 2013), produciendo un mayor índice de satisfacción en la evaluación y percepción debido a la mayor transparencia (Berrocoso y Gómez, 2014; Panadero y Jonsson, 2013). Por otro lado, desde el punto de vista del profesorado, la acción de diseñarlas muestra valores respecto a cómo se percibe el aprendizaje y permite evaluar de forma coherente y homogénea (Berrocoso y Gómez, 2014). No obstante, desarrollar rúbricas de calidad es una tarea compleja y una mala calidad de la herramienta puede afectar a su fiabilidad y a la calidad del feedback (Lovorn y Rezael, 2011).

La inclusión del alumnado en tareas de evaluación se ha observado como beneficiosa en el proceso de aprendizaje en comparación con los métodos tradicionales. Por este motivo, la aplicación de instrumentos y herramientas que fomenten su participación ha aumentado en los últimos años. En este sentido, las rúbricas se han visto efectivas en la mejora de las calificaciones, en la mejora de la autorregulación y en la percepción de autoeficacia, entre otros aspectos. Por este motivo, es de gran interés el estudio de dicho instrumento en asignaturas de diferente naturaleza con el objetivo de estudiar su impacto entre diferente tipo de alumnado.

2. Objetivos

Se estableció como objetivo principal del estudio identificar posibles diferencias en los resultados de las evaluaciones realizadas por el profesorado y el alumnado, así como analizar la fiabilidad inter-evaluador de la rúbrica diseñada.

3. Desarrollo de la innovación

3.1. Alumnado sobre el que se ha desarrollado la intervención

La intervención se llevó a cabo en el primer cuatrimestre del curso 2021-2022, en la asignatura “Biomecánica de la Actividad Física y el Deporte” del Grado de Ciencias de la Actividad Física y el Deporte, de la Universitat de València. Dicha asignatura se imparte en el tercer curso del grado. La intervención se realizó sobre la parte práctica de la asignatura. Dicha parte práctica se compone de 4 subgrupos, estando en cada subgrupo 15 estudiantes matriculados. Las sesiones prácticas fueron impartidas por 3 docentes diferentes.

La docencia práctica de la asignatura consiste en 5 sesiones. En cada sesión práctica se enseña al alumnado sobre el uso y aplicación práctica de algún instrumento de medición biomecánica. Tras cada sesión, el alumnado debe llevar a cabo de manera grupal (entre 3 y 5 personas por grupo) la realización y entrega de un trabajo escrito de 4 páginas en el que se describe la práctica y se analizan los datos obtenidos en dicha sesión. Este trabajo escrito debe poseer la siguiente estructura: objetivos del taller, material necesario, metodología, tratamiento de los datos y resultados, discusión, conclusiones y bibliografía. Los trabajos deben entregarse a través de una plataforma docente Moodle “Aula virtual” en un plazo no superior a 20 días tras la realización de la sesión práctica.

3.2. Diseño de la intervención

La intervención consistió en el desarrollo por parte de las tres docentes responsables de la práctica una rúbrica asociada a cada taller para que el alumnado la tuviese en cuenta a la hora de desarrollar el trabajo escrito, y que el profesorado la utilizase para evaluar los trabajos entregados. Además, se le pidió al alumnado que entregasen el trabajo escrito junto a su autoevaluación mediante el uso de la rúbrica. El alumnado tenía a su disposición la rúbrica en el aula virtual previamente al inicio de la misma y su cumplimentación se estableció como requisito para poder optar a la evaluación de dicho trabajo.

Aunque en total se formaron 14 grupos de trabajo en los 4 subgrupos, únicamente se han analizado para esta comunicación los datos de 10 grupos (compuestos por un total de 48 personas), ya que 3 grupos no entregaron la autoevaluación de su rúbrica, y un grupo no entregó todos los trabajos. Cada docente tuvo asignado uno o dos talleres, los cuales tuvo que realizar la evaluación mediante el uso de la rúbrica específica de cada uno de los trabajos entregados.

Por último, se escogió al azar un trabajo de tres grupos diferentes de la misma práctica. Dichos tres trabajos fueron evaluados por los tres docentes con el objetivo de conocer de cada uno de los ítems de la rúbrica la repetibilidad inter-evaluador, entendido como la concordancia en los resultados entre las personas evaluadoras. Asimismo, las tres personas que evaluaron dichos trabajos fueron las mismas que llevaron a cabo las sesiones prácticas con el alumnado.

La distribución temporal de la intervención se muestra en la Tabla 1.

Tabla 1. Distribución temporal de la intervención.

Momento	Fecha
Desarrollo de las rúbricas	Del 06/09/2021 al 26/09/2021
Realización sesiones prácticas	Del 05/10/2021 al 23/12/2021
Análisis de los resultados de la intervención	Del 20/01/2022 al 20/02/2022

3.3. Análisis estadístico

Para el análisis estadístico se utilizó el programa RStudio (versión 1.2.5033). Se asumió que los datos eran no paramétricos y se analizaron las diferencias entre las notas del alumnado y las del profesorado para cada ítem mediante el test de Wilcoxon. Además, se evaluó la consistencia entre la nota del alumnado y el profesorado mediante el análisis del Coeficiente de Correlación Intraclase (CCI). Dicho coeficiente se calculó mediante el modelo aleatorio de dos factores, de tipo acuerdo absoluto, para medidas individuales. Dichos coeficientes fueron clasificados como excelentes ($CCI \geq 0,75$), buenos ($CCI 0,74-0,60$), regulares ($CCI 0,59-0,40$), o pobres ($CCI \leq 0,39$) (Cicchetti y Sparrow, 1981). Dichos coeficientes también se calcularon para evaluar la repetibilidad entre-evaluadores en los tres trabajos analizados al azar por los tres docentes. El límite de significancia fue establecido en $p < 0,05$.

4. Resultados

4.1. Desarrollo de las rúbricas

Tras las reuniones del profesorado para el desarrollo de las rúbricas, se decidió que dichas rúbricas debían tener una valoración del 1 al 4, con 6 ítems comunes para todas las sesiones prácticas, y tres ítems deberían de adaptar su contenido específico a cada práctica. Además, se incluyó un ítem inicial en forma de

“*checklist*” para que el alumnado tuviese en cuenta que los trabajos cumplieran con la extensión máxima de 4 páginas. La Tabla 2 muestra como ejemplo la rúbrica del taller de acelerometría.

4.2. Diferencias entre la evaluación del profesorado y la del alumnado

El alumnado se autoevaluó con notas más superiores a las que le fueron otorgadas por el profesorado en todos los ítems de la rúbrica ($p < 0,01$, Figura 1). Además, los CCI mostraron en todos los ítems una concordancia pobre ($CCI < 0,4$), salvo el ítem de “Discusión y Conclusión”, donde se obtuvo un $CCI = 0,46$. La exactitud en la autoevaluación hace referencia a la diferencia entre la puntuación proporcionada por el estudiante y la dada por un evaluador, siendo un factor de relevancia para la validez de la autoevaluación.

4.3. Repetibilidad inter-evaluador

La repetibilidad inter-evaluador fue excelente para 4 ítems: diseño, referencias, redacción, y discusión y conclusión (Figura 2). La repetibilidad inter-evaluador fue buena para 2 ítems: estructura y valoración general. Los restantes 3 ítems mostraron una repetibilidad regular: descripción general, tratamiento de las variables y análisis de las variables.

Tabla 2. Ejemplo de rúbrica desarrollada para el taller de acelerometría compuesta por 6 ítems que son comunes a todas las rúbricas, y 3 ítems finales específicos a cada taller.

Checklist	El taller cumple con la extensión máxima de 4 páginas. Esta extensión incluye cualquier apartado como podría ser la portada o el índice (ambos optativos).		Subrayar la opción acertada SI NO Si se responde NO, se restarán 5 puntos de la nota final.		
	Inadecuado 1	A Mejorar 2	Adecuado 3	Excelente 4	PT
APARTADO COMÚN A TODAS LAS RÚBRICAS					
Diseño (15%)	El documento se hace arduo de visualizar, al no tener un diseño atractivo. No posee encabezados de página, y el texto está sin ningún formato.	El documento posee encabezado y pie de página con el número de página, el logo de la universidad, el título etc. Sin embargo, visualmente no es atractivo. Habría que trabajarlo más para que estuviese listo para ser presentado por un servicio comercial.	Aunque el documento posee encabezado y pie de página con el número de página, el logo de la universidad, el título etc., y se nota que están cuidados los diferentes estilos de texto, no termina de ser excelente.	El diseño del documento es atractivo, con un nivel excelente, pareciendo el documento final que podría presentar cualquier empresa/institución. Posee encabezado y pie de página con el número de página, el logo de la universidad, el título etc. Además, posee otros elementos de diseño como gráficos, esquemas, fotografías, etc...	

<p>Referencias (10%)</p>	<p>Más de 2 criterios no se cumplen:</p> <ol style="list-style-type: none"> 1. Se han utilizado al menos 3 referencias de calidad (artículos científicos y capítulos de libro). 2. Se han utilizado referencias internacionales. 3. Se ha referenciado tanto en el texto como en la sección final de bibliografía. 4. Se ha utilizado correctamente el estilo APA. 	<p>Dos de los criterios no se cumplen.</p>	<p>Uno de los criterios no se cumple.</p>	<p>Se cumplen todos los criterios</p>	
<p>Estructura (5%)</p>	<p>No se ha observado una estructura clara del documento, que permita leerlo correctamente.</p>	<p>Al documento le falta alguna de las secciones, o esta sección no está desarrollada correctamente, o posee información que debería estar en alguna de las otras secciones.</p>	<p>El documento está perfectamente estructurado con todas las secciones importantes: objetivos del taller, metodología, resultados, y conclusiones/aplicaciones prácticas, pero a alguna sección le falta un poco de información y su extensión no es adecuada.</p>	<p>El documento está perfectamente estructurado con todas las secciones importantes: objetivos del taller, metodología, resultados, y conclusiones/aplicaciones prácticas. La estructura ayuda a seguir el documento correctamente y todas tienen una extensión adecuada.</p>	
<p>Descripción general: material y metodología (10%)</p>	<p>En el documento no se describe una introducción al taller, así como el material y metodología empleada.</p>	<p>El documento presenta una breve introducción, con una descripción del material adecuada, sin embargo, es nula la existencia del procedimiento metodológico empleado.</p>	<p>El documento presenta una introducción al taller, y una descripción del material detallado (con características técnicas), acompañadas de una breve descripción metodológica.</p>	<p>El documento presenta una adecuada introducción al taller, con una detallada descripción del material (características técnicas específicas y necesarias para la práctica), así una descripción metodológica acompañada de esquemas/gráficos conceptuales.</p>	

Redacción (10%)	Cada miembro del equipo ha redactado su parte y no ha revisado/trabajado lo del resto. Esto hace que el riesgo de que alguna sección no esté bien redactada y pueda contener faltas de ortografía sea alto.	Menos de la mayoría de los componentes del equipo han revisado y trabajado la redacción del documento, lo que hace que la redacción sea muy mejorable, algunas frases no estén muy bien redactadas y se haya un alto riesgo de contener faltas de ortografía o errores gramaticales.	La mayoría los componentes del equipo han revisado y trabajado la redacción del documento, lo que hace que la redacción sea muy buena, se pueda seguir con facilidad, y se tenga una confianza mayor del 75% de que no hay faltas de ortografía o frases complejas de entender.	Todos los componentes del equipo han revisado y trabajado la redacción del documento, lo que hace que la redacción sea exquisita, se pueda seguir con facilidad, y se tenga una confianza mayor del 90% de que no hay faltas de ortografía o frases complejas de entender.	
Valoración general (10%)	Teniendo en cuenta todos los elementos de valoración de esta rúbrica, el equipo piensa que el trabajo tiene una calidad insuficiente, con un gran margen de mejora.	Teniendo en cuenta todos los elementos de valoración de esta rúbrica, el equipo piensa que el trabajo tiene una calidad suficiente y se podría mejorar.	Teniendo en cuenta todos los elementos de valoración de esta rúbrica, el equipo piensa que el trabajo tiene una calidad notable.	Teniendo en cuenta todos los elementos de valoración de esta rúbrica, el equipo piensa que el trabajo tiene una calidad excelente	
APARTADO ESPECÍFICO A LA PRÁCTICA DE ACELEROMETRÍA					
Tratamiento de las variables (10%)	El documento presenta más de uno de estos errores: -No hay una representación gráfica del análisis de las variables. -No se muestra tratamiento alguno de las variables analizadas	El documento presenta alguno de estos errores: -No hay una representación gráfica del análisis de las variables. -No se ha analizado alguna variable. -Alguna variable se ha analizado incorrectamente.	El documento ha tenido en cuenta todas las variables explicadas en la clase y se muestran gráficamente. Sin embargo, no muestra gráficamente cómo se han analizado todos los pasos, lo que suscita dudas sobre su análisis.	El documento ha tenido en cuenta todas las variables explicadas en la clase, mostrándose la formulación y/o ecuaciones para su análisis. Se muestra gráficamente cómo se han analizado todos los pasos, interpretándose con facilidad que el equipo entiende a la perfección las variables y su análisis.	

Análisis de los resultados obtenidos (20%)	No se muestran por medio de una tabla o gráfica las diferencias en las variables entre condiciones.	Se muestra por medio de una tabla o gráfica las diferencias en las variables entre condiciones. No hay interpretación de los resultados obtenidos.	Se muestra por medio de una tabla o gráfica las diferencias en las variables entre condiciones. Aunque se interpretan los resultados obtenidos, no se fundamentan en estudios previos, lo que hace que dichas explicaciones no estén correctamente sustentadas.	Se muestra por medio de una tabla o gráfica las diferencias en las variables entre condiciones. También se interpretan correctamente los resultados obtenidos fundamentándolos en estudios previos.	
Discusión y conclusión (10%)	Presenta una discusión descontextualizada con el contenido de la acelerometría.	La discusión cita las variables de acelerometría analizadas (Log. Paso/frecuencias, Máximos tibia/cabeza, Ratios tibia/cabeza, y atenuación), en el objeto de estudio durante la práctica, sin embargo, no se comentan o discuten adecuadamente.	La discusión analiza las variables de acelerometría (Log. Paso/frecuencias, Máximos tibia/cabeza, Ratios tibia/cabeza, y atenuación), estructurándose los contenidos con un hilo argumental claro y ordenado. Presenta una conclusión confusa o no relacionada con los principales aspectos analizados.	La discusión está estructurada, analizando todas las variables propuestas y relacionadas fuentes bibliográficas relacionadas con la acelerometría de la temática analizada. Presenta una conclusión clara y concisa de los resultados relacionándose con los aspectos analizados.	
<p>PUNTUACIÓN TOTAL (PT) = (SUMA DE LAS PUNTUACIONES RELATIVAS * 10) / 4</p>					

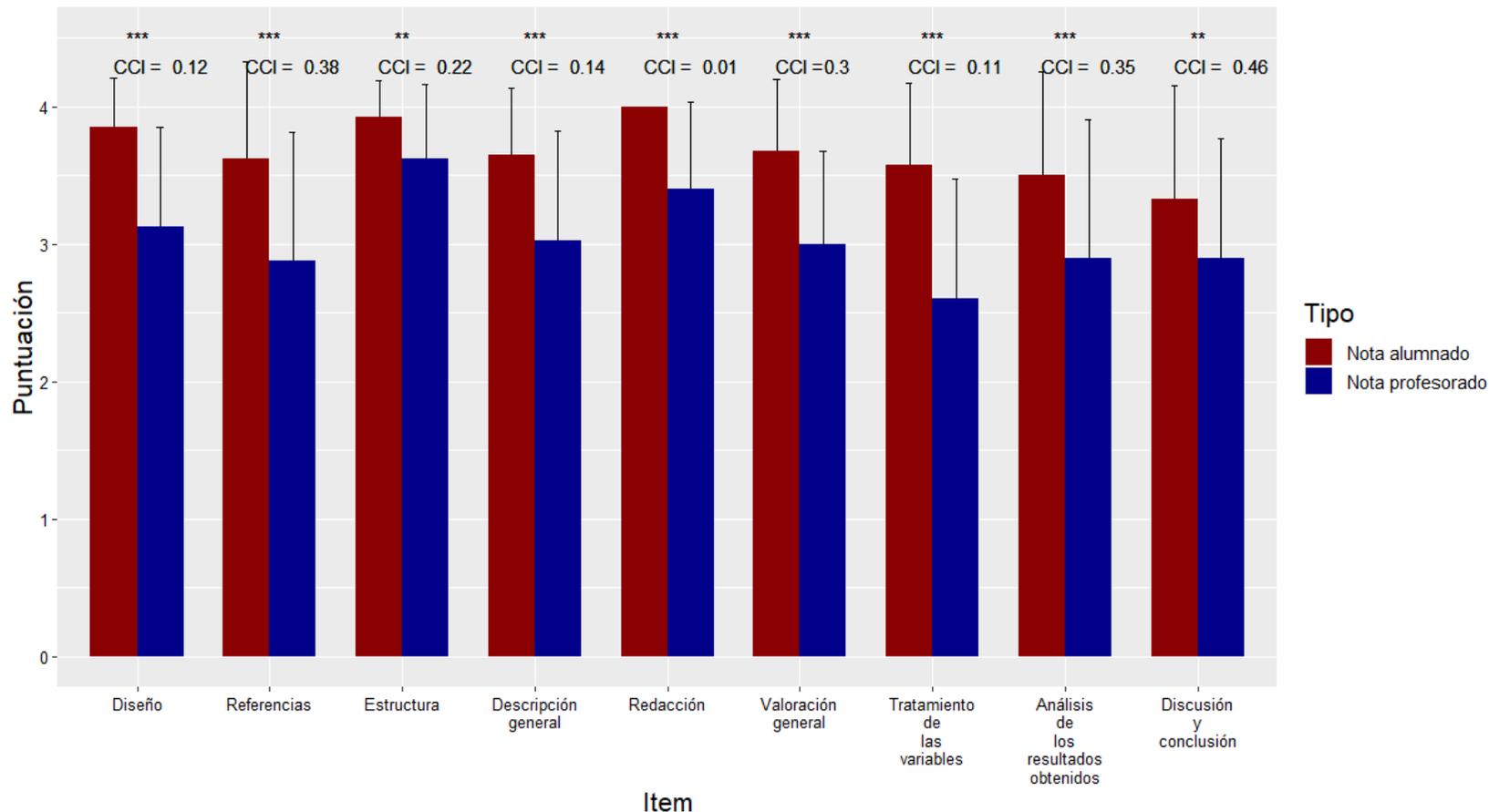


Fig. 1 Media y desviación de la puntuación otorgada en las rúbricas por el alumnado y por el profesorado para cada uno de los ítems. Las diferencias entre la nota del profesorado y el alumnado se muestran mediante símbolos (** $p < 0,01$, *** $p < 0,001$). La repetibilidad entre ambas notas se muestra mediante el Índice de Correlación Intraclass (ICC).

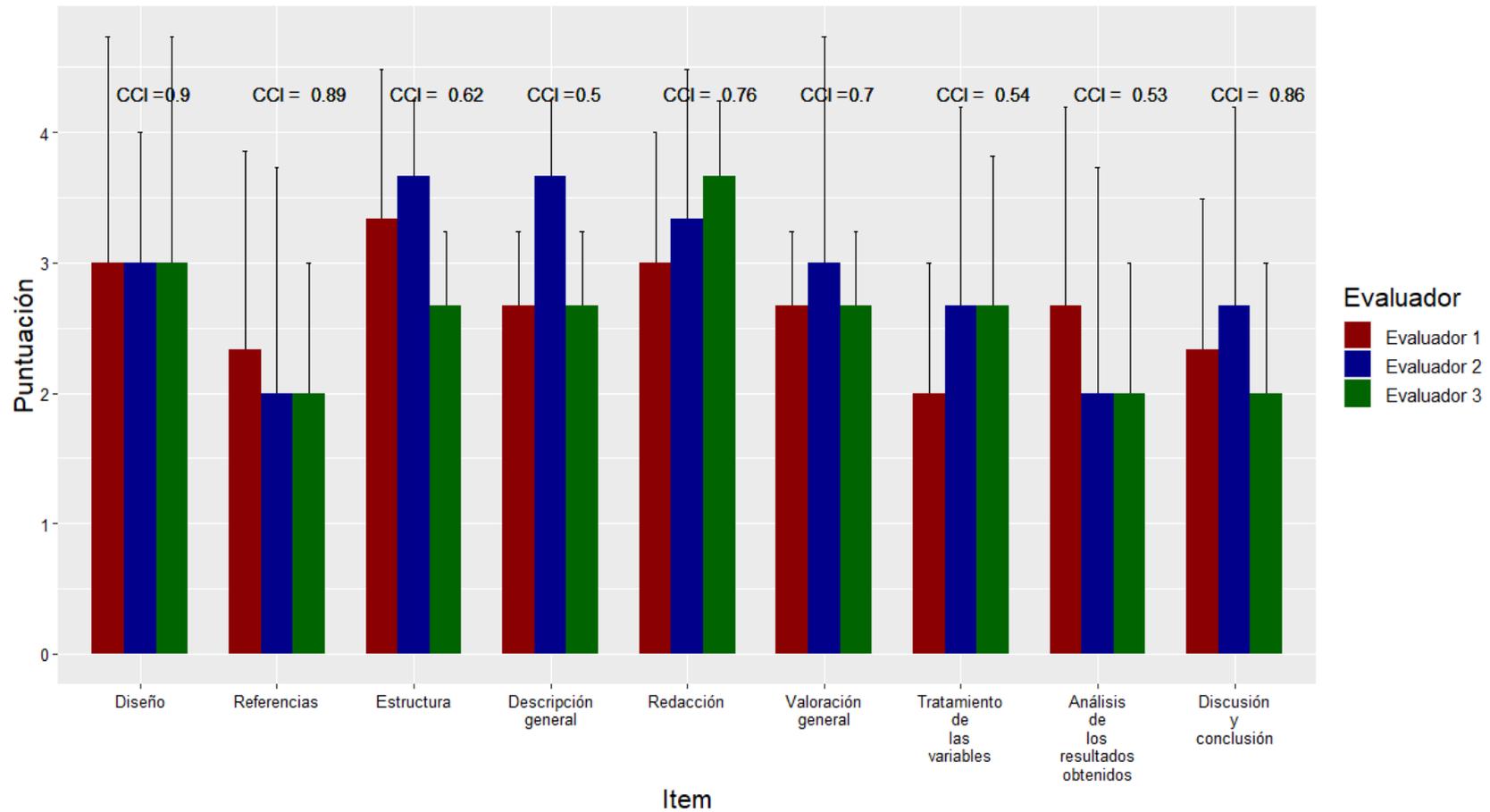


Fig. 2 Media y desviación de la puntuación otorgada en tres trabajos por los tres docentes de la asignatura. La repetibilidad entre los tres docentes se muestra mediante el Índice de Correlación Intraclass (ICC).

5. Conclusiones

El cumplimiento de los criterios de evaluación proporcionados por la rúbrica y de los diferentes niveles de logro parece no ser el único elemento involucrado en el proceso de evaluación, ya que se observan valores más elevados en todos los ítems en la evaluación del alumnado en comparación con la del profesorado. En la misma línea se encuentran los resultados obtenidos por Panadero y Romero (2013) y González y colaboradores (2017), quienes obtuvieron en su intervención calificaciones más elevadas en la auto-evaluación que en la evaluada por el profesorado. Por este motivo, sería de gran interés realizar en futuras investigaciones entrevistas al alumnado involucrado en los trabajos evaluados con el objetivo de esclarecer los motivos que provocan la disparidad en los resultados entre las personas evaluadoras. Una posible explicación a los resultados podría ser el efecto del sesgo individual del alumnado debido a que se proporciona un nivel de logro más elevado que el del profesorado.

En relación a las diferencias observadas en la repetibilidad inter-evaluador entre los ítems evaluados, se observa una repetibilidad regular en tres de los ítems evaluados (descripción general, tratamiento de las variables y análisis de las variables). Una posible explicación a la disparidad entre las personas evaluadoras podría ser que la descripción de los diferentes niveles de logro es insuficiente o imprecisa. En este sentido, la información de cada nivel de logro en dichos ítems debería ser evaluada para garantizar que son interpretados del mismo modo por parte de todas las personas implicadas.

Debido a que las rúbricas requieren de coresponsabilidad en el proceso de evaluación entre el alumnado y profesorado, la revisión de los diferentes aspectos que se han expuesto del instrumento de evaluación debería llevarse a cabo de forma conjunta con el alumnado y no solamente por parte del profesorado, lo que favorecería una mayor comprensión del método de evaluación. La revisión de las rúbricas es una tarea a realizar en clase y una propuesta didáctica podría ser realizar un *brain storming* donde se le exponga al alumnado el nombre del ítem a evaluar y a continuación expongan todos los aspectos que considerarían que deberían ser evaluados en dicho ítem. Seguidamente, la persona docente les informaría de los aspectos pre-establecidos en ese ítem así como otros nuevos a incluir tras el *brain storming*, con el objetivo de que cada una de las partes conozca las expectativas de la otra en cada uno de los ítems y niveles de logro.

6. Referencias

- Allen, D., y Tanner, K. (2006). Rubrics: Tools for making learning goals and evaluation criteria explicit for both teachers and learners. *CBE Life Sciences Education*, 5(3), 197-203. <https://doi.org/10.1187/CBE.06-06-0168/ASSET/IMAGES/LARGE/CBE0030600430002.JPEG>
- Andrade, H., y Du, Y. (2019). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research, and Evaluation*, 10(1), 3. <https://doi.org/https://doi.org/10.7275/g367-ye94>
- Berrocoso, J. V., y Gómez, A. C. (2014). El uso de e-rúbricas para la evaluación de competencias en estudiantes universitarios. Estudio sobre fiabilidad del instrumento. *REDU. Revista de Docencia Universitaria*, 12(1), 49-79. <https://doi.org/10.4995/REDU.2014.6415>
- Cicchetti, D. ., y Sparrow, S. . (1981). Developing criteria for establishing interrater reliability of specific items: Applications to assessment of adaptive behavior. *American Journal of Mental Deficiency*, 86(2), 127-137.
- Gallego, M. J., y Raposo, M. (2014). Compromiso del estudiante y percepción del proceso evaluador basado en rúbricas. *REDU. Revista de Docencia Universitaria*, 12(1), 197. <https://doi.org/10.4995/redu.2014.6423>

- González, S., Bolívar, A., Y Verano, D. (2017). Self-assessment accuracy in higher education: The influence of gender and performance of university students. *Active Learning in Higher Education*, 20(2), 101-114.
- Ibarra, M., y Rodríguez, G. (2010). Los procedimientos de evaluación como elementos de desarrollo de la función orientadora en la universidad. *REOP - Revista Española de Orientación y Psicopedagogía*, 21(2), 443-461. <https://doi.org/10.5944/REOP.VOL.21.NUM.2.2010.11558>
- Lovorn, M., y Rezael, A. (2011). Assessing the Assessment: Rubrics Training for Pre-Service and New In-Service Teachers. *Practical Assessment, Research, and Evaluation*, 16(16), 1-18.
- Panadero, E., y 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., Y Romero, M. (2013). 'To rubric or not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy. *Assessment in Education: Principles, Policy & Practice*, 21(2), 133-148.
- Real Decreto 1791/2010, de 30 de diciembre, por el que se aprueba el Estatuto del Estudiante Universitario. Boletín Oficial del Estado. Madrid, 31 de diciembre de 2010, núm. 318.
- Rodríguez, A. C., y Estrada, F. J. P. (2007). Las plantillas de evaluación (rúbrica) como instrumento para la evaluación formativa: un estudio de caso en el marco de la reforma de la enseñanza universitaria en el EEES. *Investigación en la Escuela*, 63, 77-90. <https://doi.org/10.12795/IE.2007.163.06>