



La evaluación por pares en Educación Superior

Peer evaluation in Higher Education

Lorena Atarés Huerta,¹ Juan Antonio Llorens Molina¹ y Juan A. Marin-García²

Recepción: 2020-05-27

Aceptación: 2020-09-20

Resumen

La implicación del alumnado universitario en su evaluación puede contribuir notablemente a su formación académica. Por contra, profesorado y alumnado pueden mostrarse reticentes a esta práctica, por lo que consideramos interesante realizar experiencias que testen el acuerdo de las calificaciones dadas por alumnos y profesores, así como las causas de desacuerdo. En este trabajo se realizó una experiencia de coevaluación en un curso de Química Orgánica, para comparar las calificaciones otorgadas por alumnos y profesor. Los alumnos debían realizar un póster sobre una molécula de interés agroalimentario. Se diseñó una *Rating Scale* para establecer claramente los criterios de evaluación. En la mayoría de casos no se encontraron diferencias significativas en función del evaluador ($p > 0.05$), y solamente en dos ítems de la *Rating Scale*, los alumnos otorgaron calificaciones significativamente inferiores al profesor ($p < 0.05$). En ningún caso se observaron diferencias significativas en el sentido contrario. Los grados de acuerdo entre evaluadores fueron muy dependientes del ítem de la *Rating Scale*. Nuestros resultados demuestran cómo la coevaluación puede ser una excelente herramienta para suscitar reflexiones de gran relevancia educativa, de tal modo que podríamos plantearnos en función del aspecto a evaluar, la conveniencia de asignarlo a un único evaluador (profesor o alumnado).

Palabras clave

Evaluación, evaluación del estudiante, evaluación por pares, evaluación del profesor, investigación en evaluación.

Abstract

Having our students involved in their evaluation process could remarkably contribute to their academic training. On the other hand, both teachers and students could be reluctant to this evaluation practice, which is why we consider interesting to carry out experiments testing the raters agreement in different contexts, as well as the factors influencing this agreement. In the present work we describe a coevaluation experience where the ratings given by teacher and students were compared. Our students elaborated posters about molecules of agri-food interest, in the context of their introductory course of Organic Chemistry. A Rating Scale was designed, so that the rating criteria were firmly established for both raters. In most cases, the rater had no statistically significant impact on the grades ($p > 0.05$). Only in two of the Rating Scale items the students significantly underrated the posters ($p < 0.05$). No statistically significant overrating was observed. The agreement between the raters strongly depended on the Rating Scale item considered. The discrepancies found demonstrate that coevaluation can be an excellent tool for educational reflection. Consequently, we could consider, depending on the aspect under evaluation, the suitability of having a unique rater, either the teacher or the students.

Keywords

Evaluation, student evaluation, peer evaluation, teacher appraisal, evaluation research.

¹ Universitat Politècnica de València, España.

² ROGLE- Universitat Politècnica de València, España.

Introducción

La importancia de las competencias profesionales y el papel de la evaluación en el aprendizaje son dos de los retos más relevantes en educación universitaria. Entre las competencias que demanda el entorno laboral, la comunicación efectiva es una de las más valoradas (Billington, 1997; Roncaglia *et al.*, 2008; Muñoz-Osuna *et al.*, 2016), y por eso debe trabajarse en educación superior. La comunicación escrita se percibe como un instrumento imprescindible para obtener éxito en el ámbito universitario (Gallego-Ortega *et al.*, 2013). Además, según la teoría cognitiva del aprendizaje multimedia, se aprende más profundamente si la información se presenta con palabras e imágenes que con palabras solas (Raviolo, 2019), como ocurre en la realización de pósters. Al estudiar la presentación y evaluación de pósters, Huddle (2000), destacó su potencialidad motivadora, estimulando el compromiso con el propio aprendizaje y mejorando la comprensión de los contenidos. Conyers (2003) destaca la importancia del proceso previo a la elaboración de pósters y su evaluación, señalando la importancia de la discusión previa del contenido, y la revisión crítica de sus características en función del colectivo al que van dirigidos.

Por otro lado, puede admitirse un consenso generalizado en otorgar a la evaluación un papel clave en el proceso de aprendizaje. Reyes-Cárdenas *et al.* (2019) describen el papel de la evaluación en la enseñanza de las ciencias como un proceso de construcción del conocimiento. Según Moreno-Bayardo (2019), el evaluador es un educador y su éxito debe juzgarse por lo que otros aprenden. Asimismo, Luna (2019) diferencia entre las dos funciones de la evaluación: formativa o de proceso y sumativa o de resultado. Consecuentemente, la metodología de evaluación debe cuidarse en cualquier contexto educativo (Gibbs y Habeshaw, 1989), y de hecho, el correcto diseño de su práctica puede actuar como un asistente al aprendizaje efectivo (Gibbs y Simpson, 2005).

Crear oportunidades para que el alumnado se involucre en su propio aprendizaje debe ser una prioridad para los docentes interesados en promover aprendizajes significativos (Talanquer, 2015). Asimismo, implicar al alumnado en la evaluación del aprendizaje propio y de sus compañeros puede suponer importantes ventajas, a pesar de que la docencia tradicional es reticente a aceptar esta práctica, ya que las calificaciones del profesor se consideran absolutamente válidas e indiscutibles (Orsmond *et al.*, 1996; Marin-García, J. A. 2009). Al ejercer de evaluadores, los alumnos necesitan reflexionar críticamente sobre su trabajo y el de sus compañeros, analizando aspectos positivos y a mejorar, habilidad muy útil en su futura carrera profesional, sobre todo en empresas que fomenten el trabajo en equipo (Macpherson, 1999). Marin-García, J. A. (2009) recopila una lista de ventajas de la evaluación por pares para el alumnado: aumento de la confianza en sus habilidades, mejora de la percepción sobre la calidad de su trabajo, aumento de la reflexión sobre su conducta y sus resultados, mejora de los resultados en los exámenes, mejora de la calidad y eficacia del aprendizaje, mayor responsabilidad en el aprendizaje y mayor satisfacción. La evaluación por pares es una competencia profesional en la que debemos educar a nuestros alumnos para formar profesionales reflexivos y fomentar el aprendizaje permanente (Marin-García, J. A. 2009).

Para mejorar la fiabilidad de la coevaluación, se debe establecer criterios de evaluación claros y sencillos, comunes a todos los evaluadores. En la evaluación de pósters, sin esos criterios habría riesgo de que los evaluadores fueran influenciados en exceso por el aspecto visual. Estudios previos han establecido criterios de evaluación en este tipo de actividades, a través del desarrollo y validación de *Rating Scales* (RS). Billington (1997) consideró cinco criterios para la evaluación de pósters: adecuación a las normas, claridad, diseño innovador, calidad de las fuentes y contenido. Moyo (2019) nombra los siguientes requerimientos: conforme a las normas, atractivo, claro, conciso y sin elementos innecesarios.

Objetivo

La principal inquietud que la evaluación por pares despierta en profesores y alumnos es la fiabilidad en términos de similitud con la calificación del profesor. Mejorar esta fiabilidad es clave para reducir las retenciones. El objetivo de este trabajo fue medir el grado de acuerdo de alumnado y profesor en las calificaciones de una actividad voluntaria (realización de un póster sobre una molécula de interés agroalimentario). Para ello se utilizó una *Rating Scale* diseñada expresamente.

Presentan el poster número _____, sobre la molécula _____

(Por favor, puntuad cada uno de estos ítems en una escala de 0 a 10, para vuestro poster y los otros 5 que se os indican)

Poster número...				
Sobre la molécula...				
1. considero que el poster cumple las normas que se dieron para su elaboración				
2. el poster me parece visualmente atractivo				
3. opino que el poster está equilibrado en cuanto a la proporción de texto e imágenes				
4. la distribución del texto y el tamaño de letra me parecen apropiados				
5. el poster tiene todos los contenidos que yo habría esperado para esta molécula en particular				
6. el texto me parece fácilmente comprensible y bien escrito				
7. las imágenes y gráficos me ayudan a entender mejor el contenido				
8. las fuentes de información que se han utilizado me parecen de calidad				

Gráfico 1. Rating Scale utilizada para la evaluación de los pósters

Metodología

Diseño de la *Rating Scale*

Tras la búsqueda bibliográfica pertinente, se diseñó la *Rating Scale* (RS) en base a experiencias previas similares, adaptando los ítems a los objetivos de aprendizaje que se perseguían con esta actividad. Se recopiló un número manejable de ítems (8) que recogían todos los aspectos importantes a evaluar acerca del póster: atractivo visual, adecuación a las normas, claridad, contenido y calidad de las fuentes consultadas (Gráfico 1). Dado el contexto curricular de la actividad, en la aplicación de la *Rating Scale* se priorizaron determinados criterios, previamente discutidos con el alumnado, concretamente: utilización del software recomendado para la representación de estructuras moleculares, como aspecto clave en la aplicación de las normas (ítem 1); la exigencia de citar correctamente los grupos funcionales de las moléculas orgánicas implicadas (ítem 5), así como la conveniencia de incluir manuales académicos o páginas web de química orgánica como fuentes de información y no solamente webs exclusivamente divulgativas (ítem 8). De este modo, los criterios de evaluación a seguir por el alumnado y el profesor quedaron claramente establecidos.

Aplicación de la *Rating Scale*

La experiencia se llevó a cabo en el contexto de la asignatura Fundamentos Químicos para la Ciencia y Tecnología de Alimentos, troncal de 12 créditos, estructurada en tres unidades

didácticas: Química orgánica, Termodinámica, y Cinética y equilibrios químicos. En cada unidad se llevan a cabo las correspondientes prácticas de laboratorio, coordinadas con la teoría, con la realización de actividades pre y post laboratorio.

La actividad presentada en este trabajo forma parte de la unidad I. Es importante precisar que su enfoque es eminentemente propedéutico por el enorme impacto que esta parte de la química posee en el grado, tanto desde el punto de vista académico (a continuación se cursa Bioquímica) como para la futura práctica profesional. Ello no es obstáculo para una adecuada consideración de los conocimientos previos, realizando las actividades iniciales de nivelación necesarias.

Los contenidos de la unidad se basan en estos temas principales: (1) Formulación y nomenclatura de las familias orgánicas más habituales, (2) Isomería y geometría molecular (3) Relaciones entre la estructura molecular y las propiedades físicas y la reactividad de los compuestos orgánicos y (4) Estudio descriptivo de las principales familias orgánicas, relacionadas con la ciencia y la tecnología de alimentos. Desde el punto de vista metodológico, los contenidos son desarrollados de modo participativo, mediante la realización y discusión de actividades, facilitando la información mediante múltiples recursos *online*. Cabe destacar el elevado empleo de recursos multimedia elaborados expresamente para la asignatura, así como por la importancia de la evaluación formativa mediante autoevaluaciones semanales y el empleo ocasional en el aula de herramientas como *Socrative*.

Esta experiencia se enmarca en la introducción de competencias transversales de la Universitat Politècnica de València. La actividad presentada fue un elemento imprescindible para incorporar las habilidades comunicativas y el empleo de herramientas tecnológicas. Consecuentemente los objetivos de la actividad fueron:

- Representar la estructura de una molécula de importancia agroalimentaria mediante la fórmula de esqueleto y su modelo tridimensional utilizando la aplicación informática *ChemSketch*TM.
- Especificar en dicha molécula qué grupos funcionales estudiados en el curso aparecen.
- Describir las principales propiedades físico-químicas de la molécula, haciendo referencia a aspectos de interés alimentario y de seguridad.

En la experiencia participaron voluntariamente 81 alumnos de los 90 matriculados, organizados en 27 equipos de 3, que se numeraron para facilitar la organización. Para evitar que más de un grupo elaborara su póster sobre la misma molécula, los profesores listaron 54 moléculas interesantes en base a la familiaridad de los alumnos con las mismas a través del programa de Química Orgánica y su presencia en alimentos. A cada grupo se le propuso dos moléculas para elegir una y se publicó la lista definitiva de moléculas asignadas a los grupos, así como un *Manual de orientaciones* para ayudar en la elaboración de los pósters. En este manual se indicaba expresamente la importancia de utilizar la aplicación *Chemsketch*TM para representar la molécula en el póster, junto otras premisas acerca del formato, contenido, calidad visual, no redundancia, selección de fuentes de información etc. Se añadió un apartado informando sobre los aspectos del trabajo que se valorarían y cómo (por pares y profesor), incluyendo la *Rating Scale* que se utilizaría.

Dos semanas después los alumnos entregaron un borrador de su póster para que los profesores pudieran revisarlos y asesorarlos para mejorarlos. Tras dos semanas más tuvo lugar la entrega final. Los pósters se expusieron en el hall de la Escuela durante 5 días. En este periodo todos los grupos evaluaron 5 pósters utilizando la *Rating Scale* presentada en el *Manual de orientaciones*. La asignación de los posters a evaluar se realizó de modo correlativo: los alumnos del grupo 1 debieron evaluar los pósters de los grupos 2 a 6 y así sucesivamente. De ese modo se evitaron cruces de evaluaciones y acuerdos de puntuación.

Paralelamente, el profesor de la unidad didáctica realizó su propia evaluación. La calificación final aplicada a los miembros de cada grupo fue la media entre la del profesor y el promedio de las cinco puntuaciones otorgadas por los estudiantes.

Análisis estadístico

Las evaluaciones realizadas por los alumnos y el profesor generaron una matriz de puntuaciones que se trató estadísticamente para valorar la validez de la evaluación por pares en comparación con la evaluación del profesor. En los estudios encontrados en bibliografía, el análisis estadístico (en los casos en que se describe) se limita bien a una comparación de medias y desviaciones, o bien a un análisis de correlaciones basado en el cálculo del parámetro r de Pearson. Además de lo anterior, en el presente estudio se realizaron tests de fiabilidad de escalas para hallar el Coeficiente de Correlación Intraclase (*Intraclass Correlation Coefficient*, ICC) tipo 1, utilizando SPSS 16.0. Adicionalmente se realizó un test de consistencia de los 8 ítems de la *Rating Scale* a través del parámetro α de Cronbach.

Resultados y discusión

Desarrollo general de la actividad

La actividad se desarrolló sin incidencias. Los alumnos la afrontaron con entusiasmo, apreciando el valor didáctico de la experiencia y la posibilidad de obtener mejores calificaciones globales por participar voluntariamente. No se plantearon objeciones ni directamente ni a través del foro de la asignatura. Este hecho puede interpretarse considerando que se trata de alumnado recién llegado a la Universidad, para el que cualquier actividad de aprendizaje o método de evaluación novedoso puede ejercer un efecto motivador. Las dudas, no demasiadas, estuvieron centradas en el uso del *software* para representar las estructuras moleculares.

Análisis estadístico

Comparación de medias, *underrating* y *overrating*

El Gráfico 2 muestra las calificaciones globales y para cada ítem dadas por ambos evaluadores.

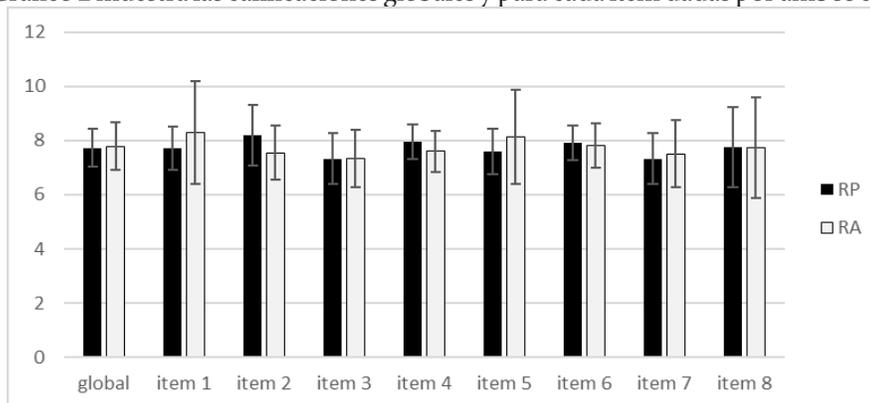


Gráfico 2. Notas globales de los pósters y en cada ítem, concedidas por el profesor (barras negras) y los estudiantes (barras grises). Las barras de error representan la desviación estándar.

El mayor motivo de reticencia frente a la evaluación por pares es el *overrating* (que los alumnos se puntúen demasiado generosamente, por encima de calificación otorgada por el profesor). Para comprobar si ocurrió en nuestro estudio, se realizó un análisis estadístico de comparación de medias en función del evaluador para todos los ítems de la RS y para la nota global. Sólo para los ítems 2 y 4 hubo diferencias significativas en función del evaluador ($p < 0.05$), y en ambos casos las calificaciones dadas por los alumnos fueron significativamente inferiores a las otorgadas por el profesor (*underrating*). En el resto de ítems, y también para la nota global, las calificaciones dadas por los alumnos no fueron significativamente diferentes de las dadas por el profesor ($p > 0.05$).

La Tabla 1 muestra el número de pósters en los que la calificación dada por los alumnos fue inferior a la del profesor (*underrating*) y el número de casos en que ocurría lo contrario

(*overrating*). Estos datos justifican las diferencias encontradas en los ítems 2 y 4, donde la distribución entre ambos está desequilibrada. En total, de las 216 calificaciones asignadas (8 ítems y 27 pósters), en 90 ocasiones el alumno puntuó más alto que el profesor, mientras que en 113 ocurrió lo contrario. En definitiva, estos resultados pueden servir para desmitificar el *overrating* y por tanto contribuyen a aportar fiabilidad a la evaluación por pares.

	Under rating	Over rating	p	r
Nota global	12	13	0.776	0.480
1. Considero que el póster cumple las normas que se dieron para su elaboración	14	13	0.101	0.351
2. El póster me parece visualmente atractivo	20	7	0.027	0.132
3. Opino que el póster está equilibrado en cuanto a la proporción de texto e imágenes	12	13	0.984	0.559
4. La distribución del texto y el tamaño de letra me parecen apropiados	15	8	0.024	0.406
5. El póster tiene todos los contenidos que yo habría esperado para esta molécula en particular	10	14	0.110	0.273
6. El texto me parece fácilmente comprensible y bien escrito	16	11	0.592	0.023
7. Las imágenes y gráficos me ayudan a entender mejor el contenido	14	11	0.397	0.565
8. Las fuentes de información que se han utilizado me parecen de calidad	12	13	0.971	0.826

Tabla 1. Número de casos de *underrating* y *overrating*, p-valor y coeficiente r de Pearson de las medidas promedio para la nota global y cada ítem de la RS.

Correlación entre las calificaciones otorgadas por alumnos y profesor

La Tabla 1 muestra el valor del coeficiente de correlación r entre las calificaciones dadas por ambos evaluadores. En el caso del ítem 4, a pesar de haberse encontrado diferencias significativas en función del evaluador, el índice de correlación encontrado es intermedio. Esto significa que hubo cierta correlación entre las evaluaciones (y por tanto una manera similar de ordenar los pósters), aunque las puntuaciones se manejaron de manera diferente por los dos evaluadores. Por otro lado, en el ítem 6 encontramos que, aunque no hubo diferencias significativas en función del evaluador ($p > 0.05$) el índice de correlación fue extremadamente bajo. Esto se debe a que realmente no hubo acuerdo entre evaluadores, pero las calificaciones dadas por los alumnos oscilaron en el mismo rango que las dadas por el profesor.

Análisis ICC

Para comprobar si los ítems son consistentes entre sí y se pueden resumir en una media global, se realizó un test de fiabilidad de las 8 puntuaciones de los alumnos y las 8 del profesor por separado. Los valores del parámetro α de Cronbach fueron 0.897 para los alumnos y 0.795 para el profesor, lo que demuestra que en ambos casos los 8 ítems tienen bastante consistencia interna. En el caso de los alumnos la consistencia es mejor, probablemente debido a un menor grado de discriminación de las notas.

El coeficiente de correlación intraclase (ICC) describe la similitud entre elementos de un grupo. Aunque puede considerarse como un parámetro de correlación, al contrario que otras medidas de correlación se basa en la comparación de datos estructurados en grupos, y no en observaciones pareadas. Este índice se utiliza para evaluar la consistencia o reproducibilidad de medidas cuantitativas hechas por diferentes observadores que miden el mismo parámetro.

Para determinar el grado de acuerdo entre evaluadores, se obtuvieron los ICC de la nota global de los pósters y en cada ítem de la RS (tabla 2).

	ICC
Nota global	0.663 (bueno)
1. Considero que el póster cumple las normas que se dieron para su elaboración	0.361 (pobre)
2. El póster me parece visualmente atractivo	0.102 (pobre)

Tabla 2. ICC de las medidas promedio para la nota global de los pósters y cada ítem de la RS.

3. Opino que el póster está equilibrado en cuanto a la proporción de texto e imágenes	0.724 (bueno)
4. La distribución del texto y el tamaño de letra me parecen apropiados	0.493 (suficiente)
5. El póster tiene todos los contenidos que yo habría esperado para esta molécula en particular	0.306 (pobre)
6. El texto me parece fácilmente comprensible y bien escrito	-0.026 (pobre)
7. Las imágenes y gráficos me ayudan a entender mejor el contenido	0.711 (bueno)
8. Las fuentes de información que se han utilizado me parecen de calidad	0.894 (excelente)

Un valor máximo de ICC de 1 revelaría un nivel de acuerdo perfecto entre los dos evaluadores. El resultado del ICC depende fuertemente del ítem considerado. La calificación global del póster tuvo un buen grado de acuerdo, así como los ítems 3, 7 y 8. El Gráfico 3 muestra las calificaciones globales de los 27 pósters, otorgadas por alumnos y profesor, así como las correspondientes a los ítems 3, 7 y 8.

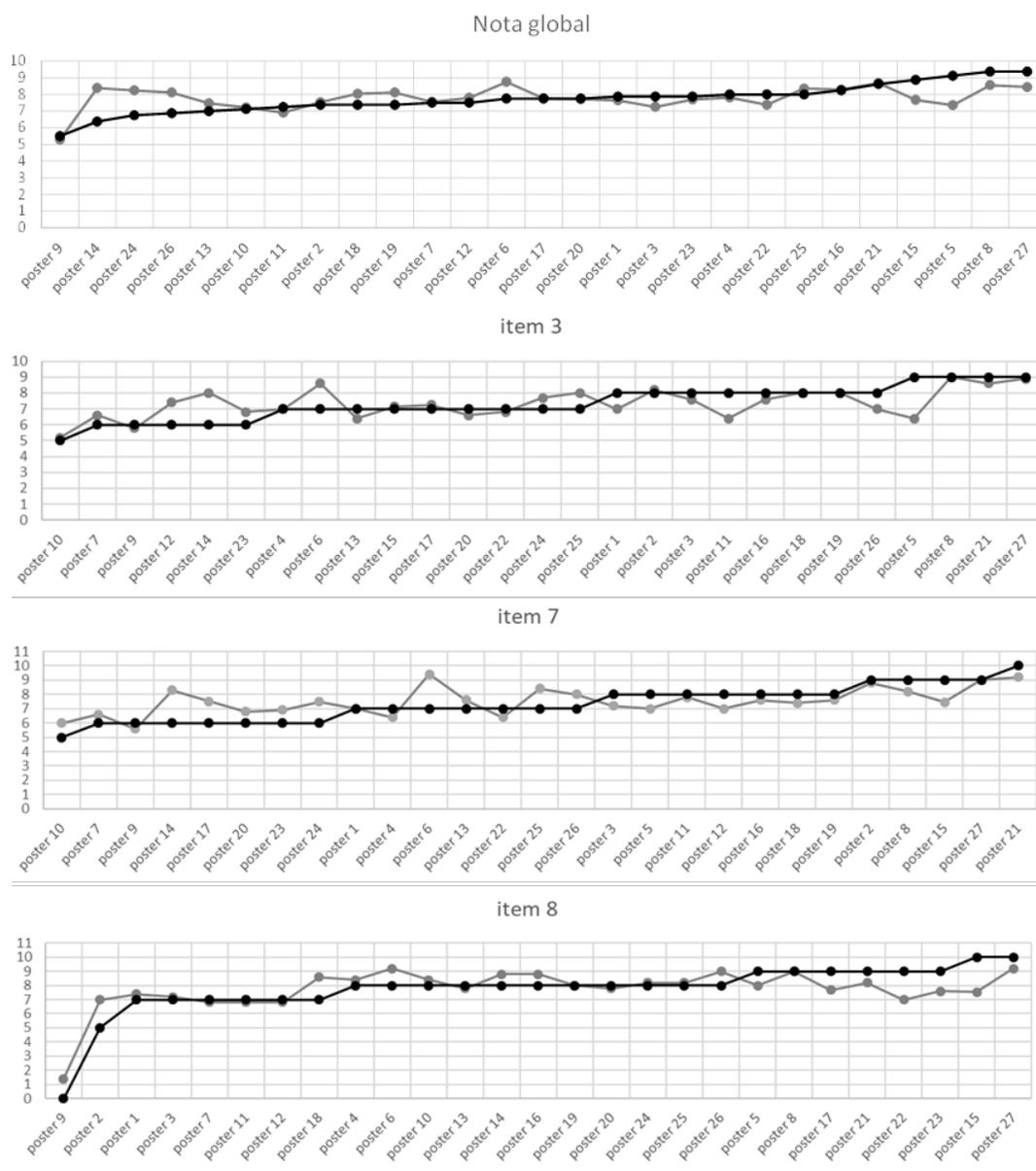


Gráfico 3. Nota global y de los ítems 3, 7 y 8 dadas por el profesor (negro) y los alumnos (gris) a los 27 pósters presentados.

Para el ítem 1, tratándose de una cuestión objetiva, podría resultar llamativo el bajo acuerdo entre ambos evaluadores. Como se muestra en los histogramas del Gráfico 4, mientras los estudiantes parecen no discriminar si los pósters cumplen o no las normas planteadas, el profesor diferencia muy claramente en función del cumplimiento de una norma en particular: la utilización de la aplicación *Chemsketch*TM para representar la molécula asignada. De este modo, la calificación del profesor se distribuyó en dos poblaciones muy diferenciadas: muchos de los alumnos fueron fieles a esta norma, y obtuvieron calificaciones más altas, mientras que otros no la siguieron, copiaron y pegaron la imagen en lugar de utilizar la aplicación, y obtuvieron una calificación menor en este ítem.

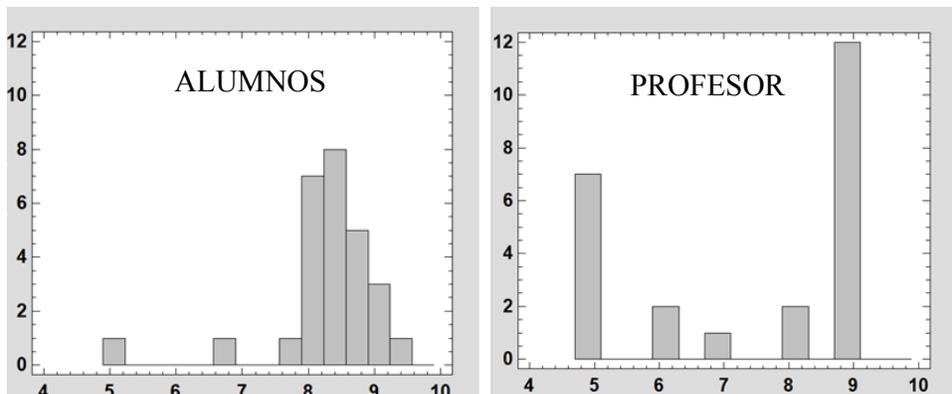


Gráfico 4. Histograma de frecuencia de notas del ítem 1 para ambos evaluadores

Los ítems 4 y 5 no presentaron un buen nivel de acuerdo, probablemente debido a su mayor carga de subjetividad, o al hecho de que pueden comprenderse de maneras diferentes en función del alumno. En concreto el ítem 5 se basa en una expectativa, que puede llegar a ser altamente dependiente del individuo.

Por último, los ítems 2 y 6 (Gráfico 5) presentaron grados de acuerdo muy pobres. En el caso del ítem 2, y como se comenta en un apartado anterior, hubo diferencias significativas entre las calificaciones dadas por ambos evaluadores ($p < 0.05$). Estos ítems hacen referencia a propiedades concretas aunque en cierto grado subjetivas de los pósters (“visualmente atractivo”) y del texto que presentan (“fácilmente comprensible” y “bien escrito”). Probablemente el bajo acuerdo entre evaluadores está relacionado con qué entienden estas dos generaciones por “visualmente atractivo”, “fácilmente comprensible” y “bien escrito”.

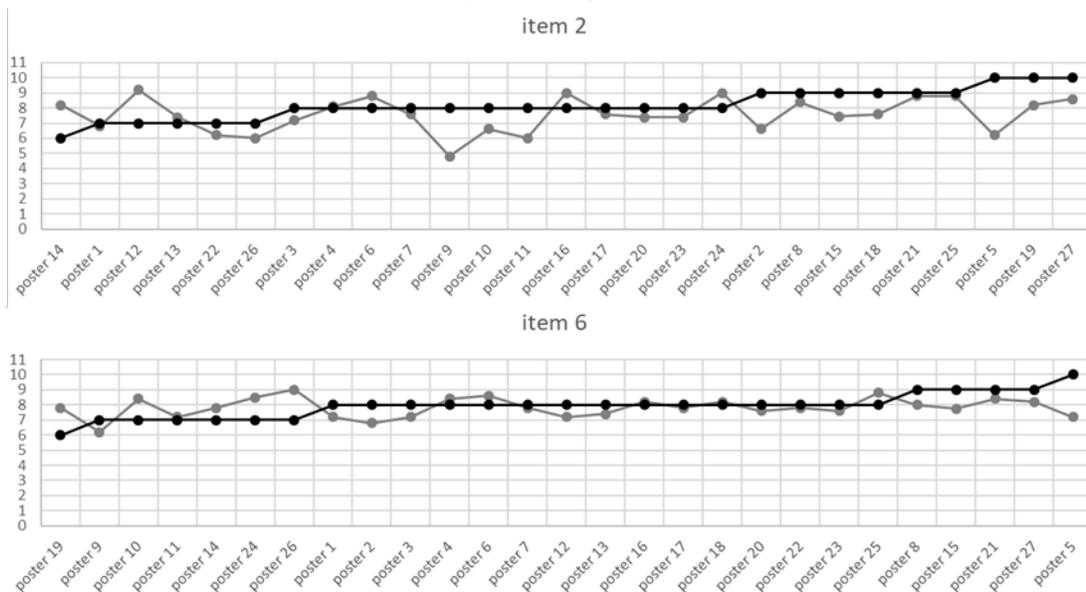


Gráfico 5. Nota de los ítems 2 y 6 dadas por el profesor (negro) y los alumnos (gris) a los 27 pósters presentados.

Conclusiones

La elaboración y evaluación de los pósters fue percibida por nuestros estudiantes como una buena experiencia de aprendizaje, puesto que les ofreció la oportunidad de desarrollar sus habilidades de comunicación gráfica en el entorno colaborativo de un grupo reducido. Además, al asumir el rol de evaluadores, se vieron en el difícil papel que normalmente corresponde exclusivamente al profesorado.

En cuanto a la metodología utilizada se comprobó la consistencia de la escala empleada. El grado de acuerdo entre las evaluaciones de alumnado y profesorado fue muy dependiente del ítem de la *Rating Scale* considerado. La posibilidad de diseñar una rúbrica lograría una descripción más detallada del grado de calidad en cada aspecto evaluado, aunque requeriría un proceso de discusión y aprendizaje por el alumnado para su utilización cuya viabilidad en el contexto concreto de la asignatura sería bastante discutible.

Las diferencias observadas entre las evaluaciones del profesorado y del alumnado pueden explicarse teniendo en cuenta la naturaleza y objetivos de la actividad. Ello se manifiesta por ejemplo en los resultados del ítem 5, en el que se reflejan las expectativas del profesorado al plantear la tarea que pueden no coincidir necesariamente con las del alumnado.

Las discrepancias observadas en determinados ítems entre la valoración del profesorado y la del alumnado muestra cómo la coevaluación puede ser una oportunidad para suscitar reflexiones y debates de gran relevancia educativa. Ello es especialmente importante cuando sus resultados afectan a la calificación y poseen por tanto consecuencias de tipo administrativo. En primer lugar cabe citar la evolución desde el punto de vista generacional de la percepción del alumnado y el profesorado acerca de aspectos tales como la calidad visual o la corrección y facilidad de comprensión de un texto escrito. Un análisis más profundo del significado que estas categorías tienen para nuestros estudiantes contribuiría a orientar mejor la realización de esta y otras tareas, así como a elaborar herramientas de evaluación más fiables. Asimismo, otra cuestión susceptible de reflexión crítica sería la conveniencia o no de delimitar aquellos aspectos en los que la coevaluación proporciona resultados convergentes y fiables, así como de aquellos otros en los que sería más adecuado asignar su evaluación al profesorado o a los estudiantes, exclusivamente. Podemos preguntarnos, por tanto, en qué aspectos de una tarea la coevaluación resulta más fiable cuando se observa un elevado grado de acuerdo. A esta pregunta solamente se puede dar respuesta desde los objetivos, contenidos y contexto de dicha tarea. En este caso, dichos aspectos serían, en primer lugar, su valoración general (síntesis de aspectos fundamentales y formales), que puede trasladarse a una calificación global de la actividad, la utilización del material gráfico en cuanto a la información que transmite y la naturaleza de las fuentes de información.

Dentro de las perspectivas abiertas por esta experiencia educativa, cabe incidir en la necesidad de ampliar este estudio refinando el diseño de la *Rating Scale* a partir de los resultados de este estudio. También cabría implicar a más cursos, así como su aplicación a otros tipos de tareas como la elaboración de videos cortos o *screencasts*.

Por último, y desde una visión más amplia de las conclusiones de este trabajo, puede plantearse una importante contradicción entre dos hechos fáciles de constatar en nuestros entornos educativos. Por una parte, la progresiva diversificación de herramientas comunicativas utilizadas, principalmente por la irrupción de las nuevas tecnologías y el acceso generalizado a los recursos multimedia. Por otra, el alumnado sigue utilizando mayoritariamente, sobre todo en las tareas de evaluación, la expresión escrita, restringida incluso al uso masivo de pruebas de opción múltiple. Desde una perspectiva bidireccional de la comunicación (lo cual es una redundancia, ya que si

no es bidireccional no es verdadera comunicación) consideramos necesaria la incorporación de herramientas mucho más diversificadas en los actos de evaluación. Actividades tales como la elaboración de un póster, una grabación en vídeo o un *screencast* pueden ser formas de expresión habituales que nuestros estudiantes empleen cuando son evaluados, más allá de su utilización complementaria para el desarrollo de ciertas competencias.

Bibliografía

- Marin-Garcia, J. A. (2009)
- Billington, H. L. (1997). Poster presentations and peer assessment: Novel forms of evaluation and assessment. *Journal of Biological Education*, 31(3), 218-220.
- Conyers, V. (2003). Posters: An assessment strategy to foster learning in nursing education. *Journal of Nursing Education*, 42(1), 38-40.
- Gallego-Ortega, J. L., García-Guzmán, A., y Rodríguez-Fuentes, A. (2013). Cómo planifican las tareas de escritura estudiantes universitarios españoles. *Revista mexicana de investigación educativa*, 18(57), 599-623.
- Gibbs, G. y Habeshaw, T. (1989). Preparing to teach. Bristol: Technical and Educational Services.
- Gibbs, G. y Simpson, C. (2005). Conditions under which assessment supports students' learning. *Learning and teaching in higher education*, (1), 3-31.
- Huddle, P. A. (2000). A poster session in organic chemistry that markedly enhanced student learning. *Journal of Chemical Education*, 77(9), 1154.
- Luna, E. (2019). Evaluación formativa del modelo educativo en instituciones de educación superior en México. *Revista mexicana de investigación educativa*, 24(83), 997-1026.
- MacPherson, K. (1999). The development of critical thinking skills in undergraduate supervisory management units: Efficacy of student peer assessment. *Assessment & Evaluation in Higher Education*, 24(3), 273-284.
- Moreno Bayardo, M. G. (2019). Temas clave de la evaluación de la educación básica. *Revista mexicana de investigación educativa*, 24(81), 633-644.
- Moyo, M. (2019). The 5 Cs for Developing an Effective Poster Presentation. *Journal of Radiology Nursing*, 38(3), 210-212.
- Muñoz-Osuna, F. O., Medina-Rivilla, A., y Guillén-Lúgigo, M. (2016). Jerarquización de competencias genéricas basadas en las percepciones de docentes universitarios. *Educación química*, 27(2), 126-132.
- Orsmond, P., Merry, S. y Reiling, K. (1996). The importance of marking criteria in the use of peer assessment. *Assessment & Evaluation in Higher Education*, 21(3), 239-250.
- Raviolo, A. (2019). Imágenes y enseñanza de la Química. Aportes de la Teoría cognitiva del aprendizaje multimedia. *Educación química*, 30(2), 114-128.
- Reyes-Cárdenas, F. D. M., Cafaggi Lemus, C. E. y Llano Lomas, M. G. (2019). Evaluación y aprendizaje basado en habilidades de pensamiento en un curso de laboratorio de química general. *Educación química*, 30(3), 79-91
- Roncaglia, D. I., Rembado, F. y Porro, S. (2008). Competencias a promover en graduados universitarios de carreras científico-tecnológicas: la visión de los empleadores. *Educación química*, 19(2), 127-132.
- Talanquer, V. (2015). La importancia de la evaluación formativa. *Educación química*, 26(3), 177-179.