

# Relación entre pruebas de elección múltiple y de ensayo en los grados de Biología Humana y de Medicina de la Universitat Pompeu Fabra

Meritxell GIRVENT, Eva BAILLÉS, Elisabeth MOYANO, Jorge PÉREZ

**Introducción.** En estudios de ciencias de la salud es muy habitual el uso de pruebas de elección múltiple y pruebas de ensayo, cada una de ellas adecuadas para evaluar diversos niveles de objetivos educativos y ambas con ciertas ventajas e inconvenientes.

**Objetivos.** Determinar la relación entre resultados en ambas pruebas en estudiantes de los grados de Biología Humana y de Medicina de la Universitat Pompeu Fabra, indagar posibles diferencias en función de la longitud de las pruebas de ensayo y conocer posibles diferencias de rendimiento académico en ambas pruebas.

**Materiales y métodos.** Se seleccionaron las calificaciones en las pruebas de elección múltiple y en las pruebas de ensayo en 12 asignaturas de los dos grados estudiados y se calcularon los coeficientes de correlación de Pearson entre ambas calificaciones. También, en cada asignatura, se compararon los resultados en ambos tipos de pruebas.

**Resultados.** En todos los casos se dieron correlaciones positivas y significativas independientemente de la longitud de las pruebas de ensayo. Así mismo, en siete asignaturas existieron diferencias significativas entre las calificaciones de ambas pruebas, siendo en cinco casos favorables a las pruebas de ensayo.

**Conclusiones.** A pesar de la alta relación entre las calificaciones en ambas pruebas, se recomienda seguir usando los dos tipos por permitir evaluar niveles cognitivos diferentes. Se destaca la necesidad de una correcta elaboración de las pruebas.

**Palabras clave.** Estudiantes de ciencias de la salud. Evaluación de estudiantes. Métodos de evaluación escritos. Pruebas de elección múltiple. Pruebas de ensayo. Rendimiento académico.

## Relationship between multiple choice and essay questions in the degrees of Human Biology and Medicine of the Pompeu Fabra University

**Introduction.** In health sciences studies, multiple choice and essay questions are habitual and each of them suitable for evaluating different levels of educational objectives and with certain advantages and disadvantages.

**Aims.** To determine the relationship between scores in both tests in students of Human Biology and Medicine degrees of the Pompeu Fabra University, to investigate possible differences depending on the number of the questions, and to know possible differences in academic achievement in both tests.

**Materials and methods.** Scores in the multiple choice questions and the essay questions tests of 12 subjects of the two grades were selected and the Pearson correlation coefficients were calculated between both scores. Also, in each subject, scores in both types of tests were compared.

**Results.** In all cases positive and significant correlations were given regardless of the number of questions of the essay tests. Likewise, in seven subjects there were significant differences between the scores of both tests being in five cases favourable to the essay tests.

**Conclusions.** Despite the high relationship between the scores in the two tests, it is recommended to continue using multiple choice questions and essay questions for allowing assessing different cognitive levels. The need for proper and fair preparation of the tests is recommended.

**Key words.** Academic achievement. Assessment. Essay questions. Health science students. Multiple choice questions. Written assessment methods.

Grupo de Investigación Educativa en Ciencias de la Salud. Departamento de Ciencias Experimentales y de la Salud. Facultad de Ciencias de la Salud y de la Vida. Universitat Pompeu Fabra. Barcelona, España.

### Correspondencia:

Dr. Jorge Pérez Sánchez. Grupo de Investigación Educativa en Ciencias de la Salud. Universitat Pompeu Fabra. Doctor Aiguader, 80. E-08003 Barcelona.

### E-mail:

jordi.perez@upf.edu

### Recibido:

13.02.19.

### Aceptado:

20.02.19.

### Conflicto de intereses:

No declarado.

### Competing interests:

None declared.

© 2019 FEM



Artículo *open access* bajo la licencia CC BY-NC-ND (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

ISSN: 2014-9832

ISSN (ed. digital): 2014-9840

## Introducción

La evaluación de los aprendizajes es un elemento determinante en la educación porque modula la forma en que los estudiantes abordan sus aprendizajes [1-3]. Tendría cinco grandes metas: determinar el grado de asunción por los estudiantes de los objetivos educativos, garantizar la futura competencia profesional de los graduados, garantizar la equidad, favorecer el aprendizaje de los estudiantes e informar a la institución sobre la marcha de los procesos docentes de las diferentes asignaturas de una carrera [4]. Así, la evaluación de los estudiantes ha de ser rigurosa y, por tanto, deben utilizarse metodologías acordes con los tipos de aprendizaje que se pretende evaluar. La comunidad educativa dispone de numerosas alternativas de evaluación para lograrlo [5-7].

En la evaluación del rendimiento académico en estudios de ciencias de la salud es muy habitual el uso de pruebas de elección múltiple y de pruebas de ensayo, de mayor o menor extensión [8-11]. Cada una de estas formas es especialmente adecuada para evaluar diversos niveles de objetivos educativos y ambas presentan determinadas ventajas, pero también ciertos inconvenientes.

Como se ha reportado [1,6,8,12,13], entre otros aspectos, las pruebas de elección múltiple serían adecuadas para evaluar resultados de conocimientos, comprensión, aplicación y análisis, pero no serían útiles para medir capacidad de síntesis y evaluación. Por el contrario, las pruebas de ensayo serían óptimas para evaluar estos dos altos niveles cognitivos. Las pruebas de elección múltiple, por el hecho de tener muchos ítems, permiten un buen muestreo representativo del contenido de un curso y, por tanto, garantizar una buena validez de contenido, lo que no ocurriría en las pruebas de ensayo.

En referencia a la elaboración y a la corrección, las pruebas de elección múltiple tienen un alto grado de dificultad y requieren un tiempo considerable, pero son muy fáciles de corregir. Contrariamente, las pruebas de ensayo serían fáciles de formular, pero difíciles y costosas de evaluar. Con respecto a la calificación de los estudiantes, las pruebas de elección múltiple serían objetivas, rápidas y fiables, mientras que las de ensayo serían más subjetivas, demoradas y menos fiables. También ambos tipos de pruebas tienen efectos diferentes sobre el aprendizaje de los alumnos, ya que las pruebas de elección múltiple ayudan a recordar, interpretar y analizar ideas de otros, especialmente del profesor, mientras que las de ensayo permiten organizar, integrar y expresar las ideas de los propios estudian-

tes. Un gran valor educativo que tienen las pruebas de elección múltiple sobre las pruebas de ensayo es que permiten un *feedback* inmediato si se dispensa la información oportuna a los alumnos.

La Facultad de Ciencias de la Salud y de la Vida de la Universitat Pompeu Fabra imparte los grados de Biología Humana, heredero de la extinta licenciatura de Biología, y de Medicina, este grado impartido de forma conjunta con la Universitat Autònoma de Barcelona. Los estudiantes de ambas titulaciones comparten buena parte de sus estudios, en concreto el 40,4% en el caso de Biología Humana y el 26,9% en el de Medicina [14]. En la Universitat Pompeu Fabra, la organización docente se organiza en trimestres y las asignaturas de las titulaciones se cursan en un único trimestre.

Desde sus inicios, la facultad dispone de un proyecto educativo donde la evaluación de los aprendizajes de los estudiantes ocupa un lugar predominante. Esta evaluación representa una tarea colectiva dirigida desde el decanato y llevada a cabo por un organismo técnico educativo (actualmente, la Unidad de Educación en Ciencias de la Salud). El proyecto colectivo de evaluación se caracteriza, entre otros aspectos, por el empleo de evaluaciones formativas a la mitad del período docente, por el uso de sesiones de evaluación conjuntas de todas las asignaturas del trimestre y por un *feedback* inmediato al acabar los exámenes [15].

Para evaluar los aprendizajes de los alumnos, en todas las materias de ambas titulaciones se utilizan pruebas de elección múltiple y, en la inmensa mayoría de ellas, también pruebas de ensayo de extensión diversa. Como ya se ha comentado, ambos tipos de pruebas son habituales en estudios de ciencias de la salud tanto en las universidades españolas como en las del resto del mundo [8-11].

La mayoría de estudios que conocemos indican una fuerte relación positiva entre los resultados obtenidos en las pruebas de elección múltiple y en las de ensayo [12,16-19]. A pesar de dicha relación, se ha argumentado que ambos tipos de pruebas incidirían en evaluar dominios cognitivos de aprendizaje diferentes [12]. También, cuando se ha comparado el rendimiento de los alumnos en los dos tipos de pruebas, se han encontrado resultados dispares, unas veces favorables a las pruebas de elección múltiple [10,16,17,20], otras a las pruebas de ensayo [9, 19], y en otras ocasiones no se han hallado diferencias [8,18,21].

El presente trabajo pretende, como objetivo principal, determinar la posible relación entre resultados en ambas pruebas de los alumnos de nuestro centro e indagar las posibles diferencias que podrían

producirse dependiendo de si las pruebas de ensayo son más o menos extensas. Un segundo objetivo sería determinar si existen diferencias de rendimiento académico de los alumnos en función de si la prueba es de elección múltiple o de ensayo.

## Materiales y métodos

### Selección de asignaturas para la realización del estudio

Para el estudio se seleccionaron asignaturas de diferentes características y pertenecientes a los grados de Biología Humana y de Medicina. En algunos casos, dichas asignaturas fueron cursadas de forma conjunta por estudiantes de ambos grados.

En primer lugar, se seleccionaron aquellas asignaturas de los grados de Biología Humana y de Medicina en las que eran responsables los autores del trabajo. En este caso, el estudio permitía contar con estudiantes de diversas promociones. En segundo lugar, se seleccionaron asignaturas de ambas titulaciones de características diferentes impartidas durante el segundo trimestre del curso 2016-2017.

Las asignaturas con datos acumulados de diversas cohortes de estudiantes fueron: Botánica (obligatoria en primer curso de Biología Humana, siete cohortes), Fisiología Vegetal (obligatoria en tercer curso de Biología Humana, siete cohortes), Psicología (obligatoria en tercer curso de Medicina, nueve cohortes), Psicología (optativa en tercer curso de Biología Humana, siete cohortes) y Nutrición (obligatoria en cuarto curso de Biología Humana, cinco cohortes, y de Medicina, seis cohortes).

Las materias cursadas en el segundo trimestre del curso 2016-2017 fueron: Zoología (obligatoria de primer curso de Biología Humana), Neurobiología (obligatoria de cuarto curso de Biología Humana), Histología (obligatoria de segundo curso de Biología Humana y de Medicina), Aparato Locomotor (obligatoria de cuarto curso de Medicina), Patología Médico-Quirúrgica V (obligatoria en quinto curso de Medicina) y Medicina Preventiva y Salud Pública (obligatoria en sexto curso de Medicina).

### Procedimiento

Se registraron las calificaciones de la primera convocatoria de los exámenes finales acreditativos de todos los alumnos de las asignaturas seleccionadas en las pruebas de elección múltiple y de ensayo. En primer lugar, se calcularon las puntuaciones medias y las desviaciones estándares en ambas pruebas y

**Tabla I.** Asignaturas seleccionadas, número de preguntas en pruebas de elección múltiple (N-PEM) y de ensayo (N-ENS) de cada una de ellas, ubicación en el currículo y tipo de asignatura. Entre paréntesis se indica el número de cohortes en cada asignatura.

	N-PEM	N-ENS	Ubicación	Tipo de asignatura
Botánica (7)	20	10	BH 1.º	Obligatoria
Fisiología Vegetal (7)	20	10	BH 2.º	Obligatoria
Psicología BH (7)	24	20	BH 3.º	Optativa
Psicología MED (9)	30	20	MED 1.º	Obligatoria
Nutrición BH (5)	20	23	BH 4.º	Obligatoria
Nutrición MED (6)	20	23	MED 4.º	Obligatoria
Zoología (1)	20	9	BH 1.º	Obligatoria
Neurobiología (1)	20	4	BH 4.º	Obligatoria
Histología (1)	24	5	BH y MED 2.º	Obligatoria
Aparato Locomotor (1)	60	4	MED 4.º	Obligatoria
Patología Médico-Quirúrgica V (1)	50	5	MED 5.º	Obligatoria
Medicina Preventiva y Salud Pública (1)	40	21	MED 6.º	Obligatoria

BH: Biología Humana; MED: Medicina.

se establecieron las posibles diferencias mediante la prueba *t* de Student. Posteriormente, se calcularon los coeficientes de correlación de Pearson entre ambas calificaciones.

## Resultados

La tabla I presenta las asignaturas seleccionadas, el número de preguntas de pruebas de elección múltiple y de pruebas de ensayo de cada una de ellas, y su ubicación dentro de los planes de estudio de los dos grados estudiados.

En la tabla II se muestra el número de estudiantes participantes en cada asignatura evaluada y la media  $\pm$  desviación estándar obtenida en ambas calificaciones. También se presenta el valor de *t* y su significación estadística en la comparación de medias y, finalmente, el valor de la correlación entre ambas calificaciones y su significación estadística.

Como puede observarse, en siete ocasiones se dieron diferencias significativas entre los resultados de las pruebas de elección múltiple y de ensayo: en cinco casos, fueron a favor de las pruebas de ensayo.

**Tabla II.** Número de estudiantes participantes en cada asignatura y media  $\pm$  desviación estándar en pruebas de elección múltiple (PEM) y ensayo (ENS). También se presenta el valor de  $t$  y su significación estadística en la comparación de medias y el valor de la correlación entre ambas calificaciones con su significación estadística.

	<i>n</i>	PEM	ENS	<i>t</i>	<i>r</i>	Ventaja
Botánica	392	5,11 $\pm$ 2,2	5,77 $\pm$ 2,2	-8,44 <sup>c</sup>	0,75 <sup>c</sup>	ENS
Fisiología Vegetal	411	5,67 $\pm$ 2,1	6,46 $\pm$ 2,0	-10,11 <sup>c</sup>	0,70 <sup>c</sup>	ENS
Psicología BH	198	8,59 $\pm$ 1,2	8,44 $\pm$ 1,3	2,42 <sup>a</sup>	0,75 <sup>c</sup>	PEM
Psicología MED	605	7,28 $\pm$ 1,6	7,47 $\pm$ 1,5	-3,60 <sup>c</sup>	0,65 <sup>c</sup>	ENS
Nutrición BH	278	6,37 $\pm$ 1,8	6,48 $\pm$ 1,5	-1,27	0,58 <sup>c</sup>	
Nutrición MED	326	6,84 $\pm$ 1,6	6,97 $\pm$ 1,4	-1,57	0,51 <sup>c</sup>	
Zoología	64	6,99 $\pm$ 1,5	7,21 $\pm$ 1,4	-1,50	0,67 <sup>c</sup>	
Neurobiología	42	5,93 $\pm$ 1,9	6,83 $\pm$ 2,4	-3,18 <sup>b</sup>	0,66 <sup>c</sup>	ENS
Histología BH-MED	110	5,20 $\pm$ 2,2	6,83 $\pm$ 2,4	-5,36 <sup>c</sup>	0,69 <sup>c</sup>	ENS
Aparato Locomotor	52	6,31 $\pm$ 1,9	6,04 $\pm$ 2,0	1,54	0,80 <sup>c</sup>	
Patología Médico-Quirúrgica V	63	7,24 $\pm$ 1,4	7,19 $\pm$ 2,1	0,30	0,73 <sup>c</sup>	
Medicina Preventiva y Salud Pública	61	7,97 $\pm$ 1,3	7,46 $\pm$ 1,6	3,42 <sup>c</sup>	0,70 <sup>c</sup>	PEM

BH: Biología Humana; MED: Medicina; Ventaja: puntuación favorable ante diferencias significativas. <sup>a</sup>  $p < 0,05$ ; <sup>b</sup>  $p < 0,01$ ; <sup>c</sup>  $p < 0,001$ .

yo, y en dos, de las pruebas de elección múltiple. Por el contrario, en cinco ocasiones no existieron diferencias de rendimiento entre ambas pruebas. En todos los casos se obtuvieron altas correlaciones positivas entre los dos tipos de pruebas, siendo siempre significativas.

## Discusión

Los resultados del estudio son muy concluyentes respecto a la relación entre el rendimiento académico de los estudiantes en las pruebas de elección múltiple y en las pruebas de ensayo. En todos los casos sin excepción la relación fue positiva y altamente significativa. Ello estaría en consonancia con la inmensa mayoría de estudios donde la relación es alta, significativa y positiva [12,16-19].

Ligado al objetivo principal, también pretendíamos determinar si la relación entre los resultados en las pruebas estudiadas podría estar mediatizada por el número de preguntas en las pruebas de ensayo. Los resultados obtenidos indicarían que la rela-

ción positiva se mantiene siempre al margen de dicho factor, ya que las altas correlaciones se encuentran en las asignaturas tanto con menor como con mayor número de preguntas de ensayo.

Si bien no era el objetivo principal del estudio, al comparar el rendimiento de los alumnos en los dos tipos de pruebas encontramos resultados dispares. De las 12 comparaciones realizadas, en siete de ellas hallamos diferencias significativas, mientras que en las otras cinco no existieron. Teóricamente, se esperarían resultados mejores en las pruebas de elección múltiple porque se relacionarían con el reconocimiento, mientras que las pruebas de ensayo se asociarían al recuerdo. En general, el reconocimiento es más efectivo que el recuerdo [22]. De hecho, la mayoría de estudios realizados confirman mejores resultados en pruebas de elección múltiple que en pruebas de ensayo [5,10,16,17,20]. A pesar de ello, en otros estudios se han presentado resultados favorables a las pruebas de ensayo [9,19] y también se han reportado trabajos donde no existían diferencias de rendimiento [8,18,21]. Nuestros resultados fueron acordes con los tres supuestos. Cuando se dieron diferencias, contrariamente a las previsiones, en cinco ocasiones los resultados fueron favorables a las pruebas de ensayo y solo en dos ocasiones a las pruebas de elección múltiple. También, en otras cinco ocasiones no se dieron diferencias de rendimiento entre las pruebas.

A pesar de la alta relación encontrada en nuestro estudio y considerando la adecuación de cada uno de los dos tipos de pruebas para evaluar los diferentes niveles de la taxonomía de Bloom [23], abogamos por la utilización tanto de las pruebas de elección múltiple como de las pruebas de ensayos en sus diferentes variantes. Este planteamiento coincidiría con la visión de numerosos autores que han realizado estudios similares [5,9,12,17,18,24-26].

Quisiéramos recalcar la necesidad de que las pruebas de ensayo, pero especialmente de elección múltiple, se elaboren de forma correcta. Sobre las pruebas de ensayo, se debería hacer un esfuerzo para usar rúbricas u otros instrumentos con los cuales reducir al máximo la subjetividad [27]. En referencia a las pruebas de elección múltiple, si bien pueden formularse para evaluar niveles altos de tipos de aprendizaje [11,13,24,28,29], con frecuencia evalúan solamente resultados de conocimientos [13,26,30,31] y en múltiples ocasiones presentan numerosas deficiencias [8,31,32].

Este hecho contrasta con la existencia de guías para una correcta elaboración de preguntas escritas que están disponibles para todo profesional de la docencia [28,29,31].

Al igual que otros autores [8,26,30,32,33], destacamos la necesidad de formación de todos los docentes sobre evaluación de los aprendizajes, en especial de aquellos que se incorporan a la enseñanza universitaria en titulaciones de ciencias de la salud.

## Bibliografía

- Gilbert JJ. Education handbook for health personnel. 6 ed. Geneva: World Health Organization; 1992.
- Wass V, Van der Vleuten C, Shatzaer J, Jones R. Assessment of clinical competence. *Lancet* 2001; 357: 945-9.
- Wass V, Archer J. Assessing learners. In Dornan T, Mann K, Scherpbier A, Spencer J, eds. *Medical education. Theory and practice*. Edinburgh: Churchill Livingstone-Elsevier; 2012. p. 229-55.
- Harden RM, Laidlaw JM. *Essential skills for a medical teacher. An introduction to teaching and learning in medicine*. Edinburgh: Elsevier; 2012.
- Schuwirth LW, Van der Vleuten CPM. Different written assessment methods: what can be said about their strengths and weakness? *Med Educ* 2004; 38: 974-9.
- Epstein RM. Assessment in medical education. *N Engl J Med* 2007; 356: 387-96.
- Nolla-Domenjó M. La evaluación en educación médica. *Principios básicos*. *Educ Med* 2009; 12: 223-9.
- Tarrant M, Knierim A, Hayes SK, Ware J. The frequency of item writing flaws in multiple-choice questions used in high stakes nursing assessments. *Nurse Educ Today* 2006; 26: 662-71.
- Bodkha P. Effectiveness of MCQ, SAQ and MEQ in assessing cognitive domain among high and low achievers. *IJRRMS* 2012; 2: 25-8.
- Aalaei S, Ahmadi MAT, Aalaei A. A comparison of multiple-choice and essay questions in the evaluation of dental students. *International Journal of Advanced Biotechnology and Research* 2016; 7: 1674-80.
- Grainger R, Osborne E, Dai W, Kenwright D. The process of developing a rubric to assess the cognitive complexity of student-generated multiple choice questions in medical education. *TAPS* 2018; 3: 19-24.
- Becker WE, Johnston C. The relationship between multiple choice and essay response questions in assessing economics understanding. *Economic Record* 1999; 75: 348-57.
- Schuwirth LW, Van der Vleuten CPM. ABC of learning and teaching in medicine. *Written assessment*. *BMJ* 2003; 326: 643-5.
- Morante N, Pérez J. Aprendiendo juntos para trabajar juntos. Comparación del rendimiento académico en estudiantes de biología humana y de medicina en la Universitat Pompeu Fabra. *FEM* 2018; 21: 255-8.
- Pérez J, Girvent M, Moyano E. La evaluación de los aprendizajes en la Facultad de Ciencias de la Salud y de la Vida de la Universidad Pompeu Fabra (UPF) de Barcelona. Una tarea colectiva. *Rev Educ Cienc Salud* 2013; 10: 127-32.
- Pepple DJ, Young LE, Carroll RG. A comparison of student performance in multiple-choice and long essay questions in the MBBS stage I physiology examination at the University of the West Indies (Mona Campus). *Adv Physiol Educ* 2010; 34: 86-9.
- Delaram M, Sharifi A. A comparison of students scores in multiple-choice and essay questions in mother and child health examination. *Future of Medical Educational Journal* 2014; 4: 15-8.
- Walke YSC, Kamat AS, Bhounsule SA. A retrospective comparative study of multiple choice questions versus short answer questions as assessment tool in evaluating the performance of the students in medical pharmacology. *Int J Basic Clin Pharmacol* 2014; 3: 1020-3.
- Adamu LH, Yusuf AH, Taura MG. Differences in multiple choice and short answer questions performance among medical and dental trainees as reflected in their anatomy course. *Journal of Contemporary Medical Education* 2018; 6: 7-16.
- Kumar P, Ghildiyal JP, Rajpoot RS, Kumar S. Comparison of performance in physiology subject assessed by short essay type questions and multiple choice questions of first year medical students. *Asian J Med Sci* 2017; 8: 82-4.
- Salih KMA, Alshehri MAA, Elfaki OA. A comparison between students' performance in multiple choice and modified essay questions in the MBBS paediatrics examination at the College of Medicine, King Khalid University, KSA. *Journal of Education and Practice* 2016; 7: 116-20.
- Myers DG. *Psychology*. 9 ed. New York: Worth Publishers; 2010.
- Bloom BS, ed. *Taxonomy of educational objectives*. New York: David McKay; 1956.
- McCoubrie P. Improving the fairness of multiple-choice questions: a literature review. *Med Teach* 2004; 26: 709-12.
- Gupta C, Jain A, D'Souza AS. Essay versus multiple-choice: a perspective from the undergraduate student point of view with its implications for examination. *Gazi Medical Journal* 2016; 27: 8-10.
- Baig M, Ali SK, Ali S, Huda N. Evaluation of multiple choice and short essay questions items in basic medical sciences. *Pak J Med Sci* 2014; 30: 3-6.
- Torrubia R, Pérez J. La difícil objetividad de las pruebas de ensayo en la evaluación del rendimiento académico. *FEM* 2005; 8: 17-21.
- Case SM, Swanson DB. *Cómo elaborar preguntas para evaluaciones escritas en el área de ciencias básicas y clínicas*. 3 ed. Philadelphia: National Board of Medical Examiners; 2006.
- Paniagua MA, Swygert KA. *Cómo elaborar preguntas para evaluaciones escritas en las áreas de ciencias básicas y clínicas*. 4 ed. Philadelphia: National Board of Medical Examiners; 2016.
- Vanderbilt AA, Feldman M, Wood IK. Assessment in undergraduate medical education: a review of course exams. *Med Educ Online* 2013; 18: 1-5.
- Palés-Argullós J, García-Barbero M. Evaluación mediante preguntas de elección múltiple (PEM). ¿Cómo elaborar correctamente el examen? In Millán Núñez-Cortés J, Palés-Argullós J, Rigual-Bonastre R, directores. *Guía para la evaluación de la práctica clínica en las facultades de medicina*. Madrid: Cátedra de Educación Médica Fundación Lilly-UCM/ Unión Editorial; 2014. p. 149-64.
- Haladyna TM, Downing SM, Rodriguez MC. A review of multiple-choice item –writing guidelines for classroom assessment. *Applied Measurement in Education* 2002; 15: 309-34.
- Rush BR, Rankin DC, White BJ. The impact of item-writing flaws and item complexity on examination item difficulty and discrimination value. *BMC Med Educ* 2016; 16: 250.