

La Predicción Del Rendimiento Académico En Alumnos De Bachillerato. Estudio De Un Caso (Centro) En España

Miguel Ángel Broc, PhD in Psychology
University of Zaragoza, Faculty of Education, Spain

doi: 10.19044/esj.2016.v12n25p111 [URL:http://dx.doi.org/10.19044/esj.2016.v12n25p111](http://dx.doi.org/10.19044/esj.2016.v12n25p111)

Abstract

This paper studies the assessment during the first year of baccalaureate in seventy-six students. In a public educational institution of Saragossa (Spain), no differences in academic achievement between men (forty) and women (thirty-six) were found, but between two main modalities (Sciences, thirty-five students, versus Humanities and social Sciences, with forty-one students), finding differences outstrip students Sciences throughout the year, equaling to end with their counterparts. This leads us to study the hypothesis of whether the ratings could be “inflated” by teachers in Humanities students. For this they were conducted by a retrospective design “*ex post facto*” an multiple lineal regression analysis, using as a dependent variable grades in the third final evaluation, and as independent, performance in the first and second assessment and all subjects, not including any other variables. The results show different prediction equations that explaining 87% and 88% of the variance in the final performance, in the two types of modalities respectively. The values predicted by the model resulting in a significant percentage of “inflation” in the ratings of 23% in the group of Sciences and 37% in the Humanities and Social Sciences. The mean scores of the subjects discussed in three different moments are analyzed and significant difference was observed in the three evaluations, emerging some evaluation patterns in the teachers. It intuits that the “rating inflation” could begin much earlier than in the pre-university entrance course.

Keywords: Assessment, baccalaureate, public education, inflated ratings, predictor variables

Resumen

En este trabajo se aborda la evaluación de alumnos de primer curso de bachillerato (LOMCE). En un centro público de enseñanza de Zaragoza capital (España), se analizan las diferencias en rendimiento académico entre cuarenta varones y treinta y seis mujeres, así como entre dos modalidades

curriculares (Ciencias, con treinta y cinco alumnos, versus Humanidades y Ciencias Sociales, con cuarenta y uno), encontrándose diferencias que aventajan a los estudiantes de Ciencias durante todo el curso, en menor número medio de suspensos, pero igualándose con sus homólogos en la tercera evaluación final de curso. Ello condujo a plantearnos la hipótesis de si las calificaciones pudieron ser “*infladas*” por el profesorado en los estudiantes de Humanidades. Para ello se realizaron, mediante un diseño retrospectivo “*ex post facto*”, análisis de regresión lineal múltiple, utilizando como variable dependiente las calificaciones en la tercera evaluación final, y como independientes, el rendimiento en la primera y segunda evaluación y todas las asignaturas, sin incluir variables de ningún otro tipo. Se utilizó preferentemente estadística no paramétrica. Los resultados arrojan ecuaciones de predicción distintas que explican el 87% y el 88% de la varianza en el rendimiento final, en las dos modalidades de bachillerato, respectivamente. Se analizan los valores pronosticados por el modelo que arroja un porcentaje relevante de “*inflación*” en las calificaciones del 23% en el grupo de Ciencias y del 37% en el de Humanidades. Se intuye que la “*inflación de las calificaciones*” podría comenzar mucho antes que en el curso previo al acceso a la universidad.

Palabras clave: Evaluaciones, bachillerato, enseñanza pública, “*inflación*” de calificaciones, variables predictoras

Introducción

El tema de la evaluación en la ESO y en el Bachillerato, es un tema a veces polémico, con ciertos interrogantes, y que en algunas ocasiones, puede convertirse en una cuestión delicada entre el profesorado. La Ley actual en vigor (LOMCE, 2013), establece que los alumnos promocionarán de primero a segundo de bachillerato cuando hayan superado las materias cursadas o tengan evaluación negativa en dos materias como máximo. El Real Decreto 1467/2007, de 2 de noviembre, por el que se establece la estructura del Bachillerato y se fijan sus enseñanzas mínimas, regula en sus artículos 12, 13, 14 y 15 los referentes, características y procesos de evaluación y los criterios de promoción y titulación del alumnado en esta etapa. Por otra parte, los criterios de evaluación deberán concretarse en las programaciones didácticas, donde también se expresarán de manera explícita y precisa los mínimos exigibles para superar las correspondientes materias, así como los criterios de calificación y los instrumentos de evaluación que aplicará el profesorado en su práctica docente (apdo. 2, p. 21646).

En la Orden de 14 de octubre de 2008, del Departamento de Educación, Cultura y Deporte, sobre la evaluación en Bachillerato en los centros docentes de la Comunidad Autónoma de Aragón se dice que “*los resultados de la*

evaluación se expresarán mediante calificaciones numéricas de cero a diez sin decimales, considerándose negativas las calificaciones inferiores a cinco. La nota media será la media aritmética de las calificaciones de todas las materias, redondeada a la centésima más próxima y en caso de equidistancia a la superior.” (p. 21646). Por su parte, cada grupo de alumnos será objeto de tres sesiones de evaluación a lo largo del curso académico, sin perjuicio de otras que establezca el proyecto curricular del centro. Se podrá hacer coincidir la última sesión con la de la evaluación final ordinaria del curso.

Aunque el tema del rendimiento académico en la ESO y Bachillerato es un tema extenso que ha llevado a muchos investigadores a intentar encontrar variables asociadas que aporten mejores explicaciones acerca del mismo, como pueden ser, por ejemplo, las relacionadas con variables psicológicas (Broc, Gil y Tejedor, 2006), Broc (2008, 2012, 2015); Miñano (2009), con variables personales y contextuales (Álvaro et ál. 1990), con variables sociales, familiares y culturales (Harter, 2012), o los estudios relacionados con los estudios de PISA y los modelos del valor añadido en educación (Martínez-Arias, Gaviria y Castro, 2009; Martínez Arias, 2009), o con las escuelas eficaces y el análisis multinivel (Ruiz de Miguel, 2009), en este trabajo nos centraremos preferentemente en la relativa discordancia que supuestamente se da entre las calificaciones reales y las pronosticadas en alumnos de bachillerato. En una línea similar, pero en niveles preuniversitarios y en pruebas de acceso a la universidad, se encuentra, entre otros, el interesante trabajo de Moreno, Sánchez y Jiménez (2014), en el que se afirma que *“...tanto en los centros privados como en los públicos las calificaciones de Bachillerato están infladas respecto a las obtenidas en selectividad, pero contradicen la hipótesis de que en los centros privados se inflan más las calificaciones que en los públicos* (p. 244). Y más adelante dice: *“...se puede aceptar aquí que entre los hombres no existen diferencias estadísticamente significativas en el inflado de las calificaciones entre los centros públicos y privados”* (p. 260), y que *“... se puede afirmar que en los centros públicos se inflan más que en los privados las calificaciones de las mujeres de cualquier modalidad de Bachillerato y las de los hombres de la modalidad de Arte”*. (p. 261).

En consonancia con esta temática también encontramos algunos estudios interesantes como el de González-Vallinas, Álvarez, Peiró y Gregori y San Fabián (2010), en el que se constataba que la nota media del bachillerato era un buen predictor de la nota media de la prueba de acceso. Una metodología utilizando una regresión lineal simple fue llevada a cabo por Gaviria (2005), Pavía y Vila (2010) con calificaciones reales y estimadas en el bachillerato, o el de Goberna, López y Pastor (1987), que utilizaron la predicción de rendimiento como criterio para el ingreso en la universidad.

En esta investigación las únicas variables fueron las calificaciones obtenidas por los alumnos en todas sus asignaturas en las tres evaluaciones del curso, así como el número de suspensos obtenidos por cada uno de ellos en esas evaluaciones. Con esto se pretende obtener un modelo de predicción del éxito o del fracaso en la evaluación final del curso, mediante técnicas de análisis de regresión múltiple. En esta línea cabría preguntarse una serie de cuestiones:

- ¿Es posible realizar una evaluación objetiva del alumnado a lo largo de todo el curso?
- ¿Existen tendencias o patrones evaluadores en el profesorado que condicionan las calificaciones de los alumnos?, ¿Y estas pueden estar mediadas por el conocido efecto Pygmalión en el aula y lo relacionado con las expectativas (Santrock, 2011)?
- ¿Hasta qué punto el esfuerzo debe incrementar una calificación si el alumno/a no llega a los objetivos o competencias mínimas?
- ¿Los mínimos curriculares de las materias del bachillerato, realmente se llevan a la práctica, o cada profesor establece sus mínimos de forma distinta o en diferentes “alturas”? Y si eso es así, ¿Quién supervisa esos mínimos, que deben adecuarse a los mínimos establecidos en el currículum oficial?

Método Muestra

Este trabajo se centró en los procesos evaluadores que tienen lugar en un centro público de enseñanza secundaria y bachillerato ubicado en una zona de clase media-baja de Zaragoza capital. Los datos fueron obtenidos en la Secretaría del Centro y a través del Sigad (Sistema de Gestión Académica y Didáctica), con el consentimiento informado, y preservando el anonimato de alumnos y profesores. El número de alumnos en la modalidad de Ciencias fue de 35, (15 varones y 20 mujeres), y en el de Humanidades y Sociales de 41 (25 varones y 16 mujeres), que hacen un total de 76 alumnos. Como el número de sujetos era bajo se utilizó preferentemente la estadística no paramétrica en el análisis de datos.

Objetivos

Utilizando un diseño retrospectivo “ex post facto” se pretendió profundizar en los objetivos siguientes: a) Comparar las puntuaciones entre los alumnos de Bachillerato en las dos modalidades principales en las tres evaluaciones del curso, y averiguar si existen diferencias estadísticamente significativas en función del sexo; b) elaborar un modelo de predicción del rendimiento final basado en los modelos de regresión lineal multivariante, a partir de las calificaciones obtenidas en la segunda evaluación; c) analizar las puntuaciones medias en todas las asignaturas del Bachillerato de las dos modalidades, y detectar cuales de ellas son estadísticamente significativas

tomadas dos a dos, entre las tres evaluaciones; d) identificar, si ese fuere el caso, patrones evaluadores emergentes en el profesorado y describir sus características principales, de cara a futuras investigaciones.

Hipótesis

a) En principio, no tendrían que aparecer diferencias estadísticamente significativas en el rendimiento en función del sexo; b) se darían diferencias significativas en el número de suspensos (o aprobados) en todas las evaluaciones en función de la modalidad de bachillerato; c) Existe una ecuación de regresión con alto valor predictivo que emerge a partir de solamente estas variables curriculares cuantitativas; d) Las diferencias entre los valores reales y los pronosticados reflejan una “*inflación*” en las calificaciones, especialmente en el bachillerato de Humanidades y Sociales, que se materializa, al final, en un mayor número de aprobados. e) Existen patrones evaluadores del profesorado más frecuentes que otros, que se pueden identificar como una variable importante a tener en cuenta en ulteriores trabajos empíricos.

Variables

En el bachillerato de Ciencias ($n = 35$) hay dieciséis variables correspondientes a las distintas asignaturas. Solamente en siete de ellas todos los alumnos del grupo ($n = 35$), están matriculados: *CC* (Cultura Científica), *EF* (Educación Física), *FI* (Filosofía), *FQ* (Física y Química), *LC* (Lengua Castellana), *MA* (Matemáticas I), *IN* (Inglés). En el resto, el número de alumnos es variable: *AN* (Anatomía, = 10), *BG* (Biología = 23), *DT* (Dibujo Técnico = 12), *RCA* (Religión Católica = 8), *FR* (Francés = 8), *TI* (Tecnología Industrial = 8), *TIC* (Tecnología de la Información y la Comunicación = 6), *EC* (Educación para la Ciudadanía = 27), y *PSI* (Psicología = 3).

En el Bachillerato de Humanidades y Ciencias Sociales ($n = 41$) hay quince variables que se corresponden con las diversas asignaturas, y todos los alumnos están matriculados en cinco materias: *CC* (Cultura Científica), *EF* (Educación Física), *FI* (Filosofía), *HMC* (Historia del Mundo Contemporáneo) e *IN* (Inglés). En las restantes, el número de alumnos varía: *ECO* (Economía = 27), *GR* (Griego I = 14), *LAT* (Latín I = 21), *MCS* (Matemáticas aplicadas a las Ciencias Sociales I = 20), *RCA* (Religión Católica = 14), *FR* (Francés = 5), *TIC* (Tecnología de la Información y la Comunicación = 15), *EC* (Educación para la Ciudadanía = 27) y *PSI* (Psicología = 21). Estas variables son ordinales y de intervalo, ya que la diferencia entre una calificación y otra es la misma.

Las tres variables restantes son el número de suspensos en la primera evaluación (*Rend1^o*), la segunda (*Rend2^o*) y la evaluación 3^a final (*Rend3^o*). Todas las variables adoptan una escala de razón, ya que son ordinales, con el

mismo intervalo y además existe un cero absoluto (no suspender ninguna materia).

Procedimiento

Después de la asistencia a las Juntas de evaluación, se descargaron del SIGAD las calificaciones de todos los alumnos/as del primer curso de Bachillerato, en las tres evaluaciones y en todas las asignaturas. Después de un largo proceso de volcado en el editor de datos del programa estadístico SPSS (*versión 22*), se llevaron a cabo los análisis correspondientes, de tipo paramétrico y no paramétrico, así como algunos análisis sobre diferencia de proporciones, estos últimos a través del programa *MedCalc* (1993-2016).

Resultados y Discusión

Un análisis de varianza del promedio de suspensos en las tres evaluaciones en función del sexo en las dos modalidades de bachillerato, no arroja diferencias estadísticamente significativas. Con 40 varones y 36 mujeres la media M de suspensos en la 1ª evaluación = 2.60 y 2.56, respectivamente ($SD = 2.28$ y 2.79 , respectivamente) y Anova ($F = .00$, $Sig = .94$). En la 2ª evaluación con una $M = 2.63$ y 2.92 ($SD = 2.08$ y 3.10 , respectivamente) y un Anova ($F = .24$, y $Sig. = .63$), y en la evaluación tercera final con una $M = 1.95$ y 2.25 ($SD = 2.37$ y 3.28 respectivamente), y Anova ($F = .21$ y $Sig. = .65$). Como los valores de la significación superan el nivel de $p < .05$, se mantiene la hipótesis nula de que no hay diferencias entre varones y mujeres en rendimiento.

Si se analizan las diferencias en rendimiento en las tres evaluaciones en función de la modalidad Ciencias versus Humanidades, puede observarse que la media de suspensos es menor en el grupo de la modalidad de Ciencias durante todo el curso académico.

TABLA I. Puntuaciones medias en el número de suspensos entre los grupos de Ciencias y Letras en las tres evaluaciones del curso

	Grupos C/L	N	Media	Desviación estándar	Media de error estándar
Rend1ª	Ciencias y Tecnología	35	1,71	2,29	,39
	Humanidades-Sociales	41	3,32	2,49	,39
Rend2ª	Ciencias y Tecnología	35	1,94	1,99	,34
	Humanidades-Sociales	41	3,46	2,86	,45
Rend3ª	Ciencias y Tecnología	35	1,51	2,32	,39
	Humanidades-Sociales	41	2,59	3,12	,49

Un contraste de medias para muestras independientes mostrará si esas diferencias son estadísticamente significativas ($p < .05$) entre ambos grupos en las tres evaluaciones (véase Tabla 2). En ella puede apreciarse que existen

diferencias significativas en rendimiento en la 1ª y 2ª evaluación, pero desaparecen en la 3ª (*Sig. bilateral* = 1.00 y .09), asumiendo y no asumiendo varianzas iguales.

A pesar de la posibilidad real de que en la última evaluación puedan desaparecer las diferencias en el número medio de suspensos entre los alumnos de los grupos de Ciencias y de Humanidades, no deja de ser extraño que se igualen en el último trimestre de curso, cuando se supone que los datos mantendrían la misma tendencia. En este sentido, existen algunos estudios como el llevado a cabo por Broc (2012), en el que se afirma que *... los alumnos de la modalidad de Ciencias aventajan a los de Letras en la mayoría de las variables metacognitivas y volitivas que se relacionan con su aprendizaje y rendimiento, aunque solamente son estadísticamente significativas a favor de los primeros las variables independientes de procesamiento de la información, estrategias de preparación de los exámenes, estrategias de aprendizaje y estudio en general, y técnicas de reducción del estrés. Además, obtienen un menor número de suspensos en todas las evaluaciones y la suma de todas las calificaciones finales es mayor* (p. 75).

TABLA II. Prueba de muestras independientes. Diferencias en rendimiento (número de suspensos) en las 3 evaluaciones según la modalidad de bachillerato (Ciencias vs Humanidades-Sociales)

		Prueba de Levene de calidad de varianzas		prueba t para la igualdad de medias						
		F	Sig.	t	gl	Sig. (bilat)	Dif.de medias	Dif. de error estándar	95% de intervalo de confianza de la diferencia	
									Inf.	Sup.
Rend1ª	Se asumen varianzas iguales	,63	,43	-2,89	74	,005	-1,60	,55	-2,70	-,50
	No se asumen varianzas iguales			-2,91	73,56	,005	-1,60	,55	-2,69	-,51
Rend2ª	Se asumen varianzas iguales	4,92	,03	-2,63	74	,010	-1,52	,57	-2,66	-,37
	No se asumen varianzas iguales			-2,71	71,34	,008	-1,52	,56	-2,63	-,40
Rend3ª	Se asumen varianzas iguales	3,43	,07	-1,66	74	,100	-1,07	,64	-2,35	,21
	No se asumen varianzas iguales			-1,70	72,58	,092	-1,07	,63	-2,32	,18

Respecto a las variables que predicen el rendimiento en la evaluación final, se podría hipotetizar que, dadas las distintas asignaturas que cursan los alumnos de Ciencias y de Humanidades, y la diferencia en rendimiento que obtienen en las evaluaciones los primeros respecto de sus homólogos, las ecuaciones de predicción del rendimiento final serán distintas. En este sentido, se llevaron a cabo análisis de regresión lineal múltiple con el fin de

obtener, los mejores subconjuntos de regresión sobre la variable dependiente número de suspensos ($Rend3^a$) en la 3ª evaluación, en los dos grupos, introduciendo las variables independientes (asignaturas), que los alumnos obtuvieron en la 2ª evaluación, además de las variables independientes etiquetadas como número de suspensos en la 1ª y en la 2ª evaluación. Los resultados de un análisis de regresión por pasos (*Stepwise*), se muestran en las tablas 3 y 4, respectivamente.

En el grupo de la modalidad de Ciencias, las variables que entran en la ecuación de predicción son, en primer lugar, el rendimiento en la 2ª evaluación (F de entrada $< .05$ y de salida 1.0), la calificación en la asignatura *TIC* y, en tercer lugar, el rendimiento en la 1ª evaluación. Todas las demás variables fueron excluidas. El resumen del modelo en el tercer paso arroja una $R = .935$, R cuadrado = $.874$, R ajustada = $.862$, un error de estimación estándar = $.863$, un cambio de cuadrado de $R = .039$, cambio en $F = 9.51$, $df1 = 1$, $df2 = 31$, $Sig.$, de cambio en $F = .00$ y un coeficiente de Durbin-Watson de 2.15. El Anova de las tres variables arroja unos valores de $F = 102.33$, 80.97 y 71.52 respectivamente, con niveles de significación $p < .00$, lo que muestra que esas variables son estadísticamente significativas. En la Tabla 3 se muestran los coeficientes no estandarizados y demás índices, a partir de los cuales se ha generado la ecuación de predicción.

TABLA III. El mejor subconjunto de regresión sobre la evaluación final en el Bachillerato de Ciencias

Modelo	Coeficientes no estandarizados		Coeficientes estandarizados Beta	t	Sig.	95,0% intervalo de confianza para B		Correlaciones			Estadísticas de colinealidad	
	B	Error estándar				Límite inferior	Límite superior	Orden cero	Parcial	Parte	Tolerancia	VIF
3 (Constante)	6,64	1,50		4,42	,00	3,58	9,70					
Rend2ª	,49	,14	,42	3,35	,00	,19	,79	,87	,51	,21	,25	3,95
SMEAN (TIC2)	-,85	,18	-,33	-4,73	,00	-1,23	-,49	-,59	-,65	-,30	,84	1,19
Rend1ª	,37	,12	,37	3,08	,00	,13	,63	,82	,48	,19	,28	3,62

a. Variable dependiente: $Rend3^a$

La mejor recta de regresión, a partir de los coeficientes no estandarizados en el tercer paso, es la siguiente: $Y (Rend3^a \text{ evaluación final}) = 6.64 + .49 * Rend2^a (n^o \text{ suspensos en la } 2^a \text{ evaluación}) - .85 * TIC (calificación en la } 2^a \text{ evaluación}) + .37 * Rend1^o (n^o \text{ suspensos en la } 1^a \text{ evaluación}) + Error$.

En el grupo de la modalidad de Humanidades y Ciencias Sociales (Tabla 4), las variables que entran en la ecuación de predicción en el tercer paso son, en primer lugar, el rendimiento en la 2ª evaluación (F de entrada $<$

.05 y de salida 1.0), la calificación en la asignatura *HMC2* (*H^a del Mundo Contemporáneo en la 2^a evaluación*) y, en tercer lugar, la puntuación real o estimada en la asignatura de *Matemáticas aplicadas a las ciencias sociales en la 2^a evaluación*. Todas las demás variables fueron excluidas. El resumen del modelo en el tercer paso arroja una $R = .94$, R cuadrado = .885, R ajustada = .876, un error de estimación estándar = 1.11, un cambio de cuadrado de $R = .02$, cambio en $F = 6.97$, $df1 = 1$, $df2 = 37$, *Sig.*, de cambio en $F = .012$ y un coeficiente de Durbin-Watson de 2.28. El Anova de las tres variables arroja unos valores de $F = 207.29$, 120.047 y 94.92 respectivamente, con niveles de significación $p < .00$, lo que muestra que esas variables son estadísticamente significativas. La mejor recta de regresión, a partir de los coeficientes no estandarizados en el tercer paso (Tabla 4), se presenta a continuación:

$$Y (\text{Rend}3^{\text{a}} \text{ evaluación final}) = -1.57 + 1.04 * \text{Rend}2^{\text{a}} (\text{n}^{\circ} \text{ Susp}) + .45 * \text{HMC2} (\text{Calificación real o estimada obtenida en H}^{\text{a}} \text{ Mundo Contemporáneo en la 2}^{\text{a}} \text{ evaluación}) - .41 * \text{SMEAN} (\text{MSC2}) \text{ o calificación real y/o estimada en Matemáticas aplicadas a las Ciencias Sociales en la 2}^{\text{a}} \text{ evaluación} + \text{Error}.$$

Puede observarse que las ecuaciones de predicción del rendimiento final en los grupos de las modalidades de bachillerato son distintas, teniendo en el de Ciencias mayor peso las dos evaluaciones previas.

Tabla IV. El mejor subconjunto de regresión sobre la evaluación final en el Bachillerato de Humanidades y Sociales

Modelo	Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.	95,0% intervalo de confianza para B		Correlaciones			Estadísticas de colinealidad	
	B	Error estándar	Beta			Límite inferior	Límite superior	Orden cero	Parcial	Parte	Tolerancia	VIF
	3 (Constante)	-1,57	1,07				-1,51	,14	-3,67	,53		
Rend2 ^a NumSusp	1,04	,09	,95	11,28	,00	,85	1,23	,92	,88	,63	,44	2,29
HMC2	,45	,14	,26	3,27	,00	,17	,72	-,51	,47	,18	,50	1,98
SMEAN (MSC2)	-,41	,15	-,21	-2,64	,01	-,72	-,09	-,69	-,39	-,15	,51	1,97

a. Variable dependiente: Rend3^aNumsusp

Volviendo a la cuestión de la diferencia en rendimiento entre las modalidades de bachillerato en la 3^a evaluación, que dejó de ser significativa, se guardaron los valores pronosticados respecto de las calificaciones reales, y la diferencia (residuos), con el fin de analizar el número de sujetos en los que las puntuaciones reales eran menores que las pronosticadas, lo que podría ser interpretado como indicio de que las puntuaciones finales pudieron ser

“*infladas*”. Los valores de influencia y los diagnósticos por caso en el SPSS nos dan una información especialmente relevante.

Esto se detecta si el número de suspensos o calificaciones reales son menores que el valor de suspensos pronosticados, así como en el coeficiente del residuo (residuos menores que 0.2 son poco problemáticos; los valores comprendidos entre 0.2 y 0.5 son arriesgados, y los valores mayores que 0.5 deberían evitarse). Por otra parte, si los residuos tipificados están normalmente distribuidos, aspecto que se asume en el análisis de regresión, cabe esperar que el 95% de ellos se encuentre en el rango [-1.96, + 1.96], y el 99% en el rango [-3, +3]. En la práctica, los casos con residuos muy grandes o muy pequeños deben ser examinados para averiguar si las puntuaciones que tienen asignadas son o no correctas, por lo que conviene estudiar esos casos para indagar si difieren de algún modo y de forma sistemática del resto de los casos. De hecho, estos casos podrían indicar el alumno/a cuya puntuación se ha inflado o devaluado. En el grupo de la modalidad de Humanidades y Sociales el número de casos inflados (cuyo número de suspensos pronosticados era mayor) fue de 13 (33%). En diecinueve casos (45%), el pronóstico coincidía con los datos reales. Mientras que en el bachillerato de Ciencias el número de casos inflados fue de 9 (26%), claramente menor que en el de Humanidades-Sociales, mientras que el modelo coincidía en el pronóstico de 18 casos, (51%).

TABLA V. Diagnósticos por casos^a en las dos modalidades de bachillerato.

Bato Ciencias (n = 35)	Número sujeto	Nº Susp. 3ª eval.	Valores Pronosticados	Residuo (Diferencia)
	9	1	2,41	-1,41
	11	4	4,64	-,64
	18	2	3,35	-1,35
	23	1	2,50	-1,50
	24	0	,89	-,89
	26	1	1,65	-,65
	31	1	2,41	-1,41
	35	0	1,27	-1,27
Bato Humanidades y Sociales (n = 41)				
	3	0	,65	-,65
	9	3	3,60	-,60
	10	2	3,23	-1,23
	12	1	2,19	-1,19
	17	0	2,34	-2,34
	19	3	3,42	-,42
	24	1	2,04	-1,04
	25	1	2,59	-1,59
	27	1	1,30	-,30
	28	4	5,91	-1,91
	29	6	6,96	-,96

34	3	4,42	-1,42
35	0	1,15	-1,15
39	2	3,38	-1,38
41	1	1,74	-,74

a. Variable dependiente: Rend3ªNumsusp

Lo que significa que alrededor del 23% en el bachillerato de Ciencias y en torno al 37% en el de Humanidades y Sociales tuvieron un número de suspensos menor que el pronosticado. El hecho de aprobar más asignaturas que las esperadas podría deberse a factores de suerte en los exámenes, a un aumento del esfuerzo que es sensible al profesorado, aunque esto no implique la superación real de los objetivos mínimos de la asignatura, pero que es reconocido “*inflando*” la calificación final, como también a estrategias de “*copia*” en los exámenes, a veces declarado por parte de muchos alumnos. Aunque también es probable un aumento del esfuerzo y del estudio a última hora, esto no parece implicar una asimilación real por parte de los alumnos, ya que algunos estudios parecen indicar que el rendimiento y la velocidad académica parece llevar una velocidad constante desde principio hasta el final de curso por parte de los alumnos, distinta en cada uno, pero que se suele mantener relativamente uniforme (Borkowski y Thorpe, 1994). También podría apelarse a motivos de “*currículum oculto*” en el sentido de que mantener la plantilla o el cupo de profesorado en un Centro implica que haya grupos suficientes en 2º de bachillerato, y ello conlleva necesariamente que haya alumnos que promocionen.

Esta cuestión nos lleva necesariamente a hipotetizar la emergencia de determinados patrones evaluadores del profesorado que permitirían “*deslizar*” las calificaciones individuales o de grupo hacia una determinada dirección. Para ello utilizamos la técnica de *Varias Muestras Relacionadas* (pruebas no paramétricas), de Friedman para averiguar si hay diferencias estadísticamente significativas, y la de Wilcoxon para dos muestras relacionadas, ya que el número de sujetos es pequeño y la distribución de algunos datos no se ajusta a la curva normal, con objeto de analizar qué variables difieren entre sí, aplicando la corrección de Bonferroni para controlar la tasa de error, es decir, para controlar la probabilidad de cometer errores de tipo I. Con tres variables es posible hacer tres comparaciones dos a dos: 1-2, 1-3, 2-3. (Véase Tabla 6).

TABLA VI. Estadísticos descriptivos y comparaciones de las puntuaciones medias en las tres evaluaciones tomadas dos a dos de las asignaturas del bachillerato de Ciencias (n = 35) y en el de Humanidades y Sociales (n= 41). (Prueba de Wilcoxon)

Ciencias y Tecno.	1ª ev.		2ª ev.		3ª ev.		1-2	2-3	1-3
	Min	Max	Min	Max	Min	Max			
	<i>M y</i>	<i>SD</i>	<i>M y</i>	<i>SD</i>	<i>M y</i>	<i>SD</i>			

Rend-N° Susp.	1.71 2.29	0-8	1.94 1.99	0-7	1.51 2.32	0-10	Z= -1.17 Sig.=.24	Z= - 2.15 Sig.= .03*	Z= - 1.25 Sig.= .21
AN (n = 10)	7.90 1.10	6-9	9.10 0.74	8-10	8.90 .88	8-10	Z= -2.40 Sig.= .01*	Z= - .82 Sig.= .41	Z= - 2.64 Sig.= .00*
BG (n = 23)	6.87 2.26	2-10	7.17 2.69	1-10	7.22 2.56	1-10	Z= -1.54 Sig.= .12	Z= - .33 Sig.= .74	Z= - 2.00 Sig.= .04*
CC (n = 35)	9.17 .78	7-10	8.23 1.31	5-10	8.60 1.24	4-10	Z= -4.07 Sig.= .00*	Z= - 2.98 Sig.= .00*	Z= - 3.39 Sig.= .00*
DT (n = 12)	6.75 2.14	3-10	7.25 2.05	3-10	7.08 1.97	4-10	Z= -1.37 Sig.= .17	Z= - 1.00 Sig.= .32	Z= - 1.04 Sig.= .30
EF (n = 35)	6.74 1.72	4-10	6.91 1.29	5-9	7.77 1.39	4-10	Z= -1.50 Sig.= .13	Z= - 4.02 Sig.= .00*	Z= - 3.83 Sig.= .00*
FI (n = 35)	6.26 2.40	2-10	7.06 2.61	2-10	6.9 2.28	1-10	Z= -2.77 Sig.= .00*	Z= - .65 Sig.= .52	Z= - 3.10 Sig.= .00*
FQ (n = 35)	5.14 2.90	1-10	4.23 2.49	1-10	4.63 2.48	1-10	Z= -2.58 Sig.= .01*	Z= - 1.65 Sig.= .09	Z= - 2.68 Sig.= .00*
LC (n = 35)	5.71 2.05	2-10	6.17 2.30	2-10	6.11 2.40	1-10	Z= -2.49 Sig.= .01*	Z= - .44 Sig.= .66	Z= - 2.50 Sig.= .01*
MAT (n = 35)	5.54 2.40	1-10	4.43 2.45	1-9	6.00 2.77	1-10	Z= -4.20 Sig.= .00*	Z= - 4.77 Sig.= .00*	Z= - 2.09 Sig.= .03*
IN (n = 35)	7.14 2.25	1-10	7.43 1.82	2-10	7.51 2.05	2-10	Z= -1.11 Sig.= .27	Z= - 1.07 Sig.= .28	Z= - 2.35 Sig.= .02*
RCA (n = 8)	9.13 .99	8-10	9.63 .52	9-10	9.75 .46	9-10	Z= -1.63 Sig.= .10	Z= - 1.00 Sig.= .32	Z= - 1.89 Sig.= .06
FR (n = 8)	7.25 1.83	5-10	8.13 1.55	6-10	8.00 2.00	5-10	Z= -2.33 Sig.= .02	Z= - .45 Sig.= .65	Z= - 2.12 Sig.= .03*
TI (n = 8)	5.38 1.60	3-8	8.25 1.49	6-10	7.13 1.55	5-9	Z= -2.37 Sig.= .02*	Z= - 2.06 Sig.= .04*	Z= - 2.23 Sig.= .03*

TIC (n = 6)	7.67 1.37	6-10	7.83 2.32	4-10	7.50 2.81	4-10	Z= -.27 Sig.= .78	Z= - .45 Sig.= .65	Z= - .38 Sig.= .07
EC (n = 27)	8.11 1.71	1-10	8.65 1.21	5-10	8.67 1.62	4-10	Z= -1.85 Sig.= .18	Z= - .85 Sig.= .40	Z= - 2.16 Sig.= .03*
PSI (n = 3)	4.33 2.08	2-6	5.33 2.52	3-8	5.00 2.00	3-7	Z= -1.34 Sig.= .18	Z= - 1.00 Sig.= .32	Z= - 1.41 Sig.= .16
Humanid. y Sociales	<i>1ª ev</i> <i>M y</i> <i>SD</i>	Min Max	<i>2ª ev</i> <i>M y</i> <i>SD</i>	Min Max	<i>3ª ev</i> <i>M y</i> <i>SD</i>	Min Max	1-2	2-3	1-3
Rend-Nº Susp.	3.32 2.49	0-9	3.46 2.86	0-10	2.59 3.14	0-10	Z= -.49 Sig.= .62	Z= - 3.70 Sig.= .00*	Z= - 2.09 Sig.= .03*
CC (n = 41)	4.49 1.82	1-9	5.63 2.02	1-10	5.56 1.82	1-10	Z= -3.51 Sig.= .00*	Z= - .34 Sig.= .73	Z= - 3.59 Sig.= .00*
ECO (n = 27)	5.48 1.63	1-9	4.93 2.64	1-10	5.30 2.43	1-10	Z= -1.76 Sig.= .08	Z= - 1.32 Sig.= .19	Z= - .23 Sig.= .81
EF (n = 41)	6.76 1.88	4-10	6.00 1.70	2-9	6.59 2.02	2-10	Z= -3.54 Sig.= .00*	Z= - 3.73 Sig.= .00*	Z= - 1.07 Sig.= .28
FI (n = 41)	4.71 2.00	2-10	4.41 2.19	1-10	4.85 2.10	1-10	Z= -.79 Sig.= .43	Z= - 2.03 Sig.= .04*	Z= - .78 Sig.= .43
GR (n = 14)	6.36 1.59	5-9	5.0 2.48	1-9	6.64 2.24	1-10	Z= -2.26 Sig.= .02*	Z= - 2.46 Sig.= .01*	Z= - .78 Sig.= .43
HMC (n = 41)	4.56 1.59	1-9	4.80 1.81	1-10	5.63 2.23	1-10	Z= -1.18 Sig.= .24	Z= - 3.12 Sig.= .00*	Z= - 4.08 Sig.= .00*
LAT (n = 21)	6.29 1.38	4-9	6.29 1.98	2-10	7.05 1.96	1-10	Z= -.14 Sig.= .88	Z= - 2.62 Sig.= .00*	Z= - 1.81 Sig.= .07
LC (n = 41)	4.32 1.66	1-8	4.78 1.59	1-9	4.73 2.13	1-10	Z= -2.74 Sig.= .00*	Z= - .31 Sig.= .75	Z= - 2.09 Sig.= .04*
MCS (n = 20)	4.50 2.18	1-8	3.90 2.29	1-8	4.45 2.44	1-9	Z= -1.95 Sig.= .05	Z= - 2.65 Sig.= .00*	Z= - .17 Sig.= .87

IN (n = 20)	4.41 2.31	1-9	3.85 2.14	1-9	4.10 2.46	1-10	Z= -4.26 Sig.= .00*	Z= - 1.77 Sig.= .07	Z= - 2.13 Sig.= .03*
RCA (n = 14)	7.29 1.64	5-10	7.86 1.61	3-10	8.64 2.27	1-10	Z= -1.58 Sig.= .11	Z= - 2.04 Sig.= .04*	Z= - 2.27 Sig.= .02*
FR (n = 5)	7.00 1.58	5-9	7.60 1.34	6-9	7.40 2.30	5-10	Z= -1.73 Sig.= .08	Z= - .38 Sig.= .70	Z= - .82 Sig.= .41
TIC (n = 16)	7.06 .85	5-8	7.20 1.52	5-10	7.60 1.40	4-10	Z= -.73 Sig.= .46	Z= - 1.73 Sig.= .08	Z= - 2.16 Sig.= .03*
EC (n = 27)	8.52 .85	7-10	7.48 2.14	1-10	7.56 2.00	3-10	Z= -2.64 Sig.= .00*	Z= - .53 Sig.= .59	Z= - 2.43 Sig.= .01*
PSI (n = 21)	4.33 1.43	1-7	4.33 1.77	1-7	4.10 1.51	1-6	Z= 0.00 Sig.= 1.00	Z= - 1.51 Sig.= .13	Z= - 1.03 Sig.= .30

(*) Las diferencias son estadísticamente significativas $p < .05$

A partir del análisis reflexivo de la tabla anterior emergerían, al menos, cuatro patrones evaluadores relevantes del profesorado, basados en las puntuaciones medias otorgadas en su asignatura a lo largo de las tres evaluaciones del curso.

- Patrón A-ascendente lineal (/), 3 de 31 asignaturas (9,67%). Comenzando con un promedio determinado de calificaciones en la 1ª evaluación, lo aumentan en la 2ª y también en la evaluación final.
- Patrón B-lineal constante (---) (22,58%, 7 asignaturas), uniforme constante en promedio de calificaciones: Profesores cuyas medias al comienzo son similares y mantienen esa tendencia a lo largo del curso, tanto en la 2ª como en la 3ª evaluación. No se constatan diferencias estadísticamente significativas entre los tres momentos temporales ni se aprecian discontinuidades.
- Patrón C1-función logarítmica, cuadrática o cúbica en forma de U (41,93%, 13 asignaturas). A partir de un determinado promedio de calificaciones en la 1ª evaluación, disminuye en la 2ª y vuelve a ascender en la 3ª. (v). Se constatan diferencias estadísticamente significativas entre la 1ª y 2ª evaluación o entre la 2ª y la 3ª, pero no entre la 1ª y la 3ª.
- Patrón C2-función logarítmica, cuadrática o cúbica en forma de U invertida (25,80%, 8 asignaturas). Comenzando en un determinado promedio de calificaciones, este aumenta en la 2ª evaluación pero finalmente vuelve a descender, siendo similar a la 1ª evaluación, sin diferencias significativas. (^)

Patrón complementario al C1. Suelen constatarse diferencias estadísticamente significativas entre la 1ª y 2ª evaluación o incluso, entre la 2ª y la 3ª, pero no entre la 1ª y la 3ª.

- Patrón D-descendente lineal (\) 0 casos (0%). Serían profesores con unas puntuaciones medias en su asignatura en la 1ª evaluación que van descendiendo progresivamente en la 2ª y la 3ª, es decir, que acaban el curso otorgando más bajas calificaciones que al principio. Este patrón es prácticamente inexistente.

Un análisis más pormenorizado de los datos, que excedería los límites de este trabajo, permitiría comparar los porcentajes de suspensos de cada profesor en cada asignatura a lo largo del curso, lo que nos llevaría a vislumbrar empíricamente que esos porcentajes no son los mismos en las diferentes materias, que van cambiando a lo largo de las evaluaciones, con diferencias estadísticamente significativas y que la dificultad de superación de las mismas es diferente desde principio de curso.

Conclusiones y prospectiva

En esta investigación ha podido constatarse que no existe una diferencia significativa entre varones y mujeres en rendimiento académico, operativizado como promedio de suspensos, pero se han constatado diferencias significativas entre las dos modalidades de bachillerato, menores en el grupo de Ciencias, pero que se igualan a final de curso con el de Humanidades. Este hecho se interpretó como una supuesta “*inflación*” en el sentido de que los valores reales en las calificaciones finales eran menores que los pronosticados, observando que este fenómeno ocurre en ambas modalidades pero que en la de Humanidades es mayor (37% frente a 23% en Ciencias) y se sospecha que la inflación de calificaciones comienza mucho antes que en 2º curso de bachillerato, tal vez desde los primeros cursos de la ESO, en base a observaciones de comportamientos del profesorado en las diversas juntas de evaluación, en las que no vamos a centrarnos ahora. También se ha mostrado que las ecuaciones de predicción del rendimiento son distintas en las dos modalidades, y que los porcentajes de varianza explicada son elevados en ambos modelos (87% en Ciencias frente a 88% en Humanidades), por lo que esta metodología puede servir a los centros en sus prácticas evaluadoras futuras, utilizando mayores muestras de sujetos y de centros y aconsejando la utilización de metodologías como el MLwin o el análisis multinivel (Gaviria, 2005, Gaviria y Castro, 2005; Martín, Martínez Arias, Marchesi y Pérez, 2008), con el objeto de si estos resultados se repiten en otros centros educativos. Por otra parte, se ha constatado que las tendencias en las puntuaciones medias de las calificaciones en las distintas evaluaciones son muy diferentes, aunque emergen ciertos patrones evaluadores que son más frecuentes que otros. Esto conduce a preguntarse si el profesorado es

consciente de sus propios patrones de evaluación y qué efecto pueden tener en el rendimiento de los alumnos.

En principio, los patrones lineales ascendentes y los uniformes parecerían más coherentes, debido a que, en primer lugar, sería de esperar que a medida que avanzara el curso los alumnos se fueran adaptando a la forma de evaluar del profesor y a la forma de estudiar la asignatura. El profesor iría observando, por su parte, progresos en la materia de forma progresiva y su práctica docente se traduciría en un mayor número de aprobados o en un menor número de suspensos a final de curso, ya que también se adapta a la situación. Los patrones uniformes serían más resistentes al cambio, en el sentido de que los profesores serían menos flexibles y tendrían las ideas más estables respecto a los objetivos mínimos de la materia, siendo poco condescendientes en los casos dudosos. Por otra parte, esto también se ajustaría a una concepción de los estudiantes relativamente constante en el sentido de que algunos llegarían a tomar conciencia de que no pueden mejorar en una materia determinada por la falta de conceptos previos bien porque no tienen tiempo y “abandonan” otras materias para septiembre, etc., o porque se sienten desmotivados con ciertas asignaturas, y prefieren optar por otras cuyas expectativas de aprobar son mayores.

Algunos patrones como los lineales descendentes (en número de suspensos), no se han tenido en cuenta. No tiene sentido pedagógico, por ejemplo, que un profesor comience el curso suspendiendo un porcentaje determinado de alumnos, y que ese porcentaje sea cada vez mayor en la segunda y en la tercera evaluación y que sus calificaciones sean cada vez más bajas.

Los patrones C1 y C2 en los que se dan diferencias estadísticamente significativas entre la primera y la segunda evaluación (de forma ascendente o descendente, y entre ésta y la tercera, siendo o no significativas entre la primera y la última, parecen más susceptibles de análisis, en el sentido de que se aproximan a distribuciones logarítmicas, cuadráticas o cúbicas, mejor que a las lineales, lo que hace reflexionar acerca de las mismas en las siguientes direcciones: a) El profesorado no tiene claros los criterios de evaluación o calificación y, a partir de un momento dado, intenta autorregularse, aumentando o disminuyendo el porcentaje de los mismos en base a algún tipo de creencia implícita sobre el número de alumnos que él/ella cree que deberían de aprobar o suspender su asignatura a final de curso, provocando, en consecuencia, un reajuste en la siguiente evaluación; b) en caso de ser conscientes de dicho patrón, esto estaría también insertado en su sistema de creencias implícitas acerca de lo que mejor les va a los alumnos o a lo que mejor ha funcionado en su práctica docente como, por ejemplo, frases del tipo “*Hay que subir el nivel de exigencia en la 2ª evaluación para que los alumnos estudien más, y así vayan más sobrados en la tercera*”, u otras del tipo “*Si*

apruebas más en la segunda evaluación los alumnos se relajan y luego acaban por suspender más en la tercera”; c) un efecto Pygmalión de expectativas positivas sesgadas hacia determinados alumnos desde comienzos de curso que podría “deslizar” la evaluación hacia el aprobado por parte del profesor, aún en contra de evidencias contrarias que se dieran en aquellos; d) Finalmente, necesidades organizativas relacionadas con el mantenimiento o la supresión de determinadas plazas de profesorado en el Centro, podría igualmente “deslizar” la evaluación positiva hacia cierto número de alumnos con el fin de mantener un número mínimo de ellos en las asignaturas que continúan en el curso siguiente. e) Es posible que algunos profesores coloquen los mínimos por debajo de los mínimos, con lo cual se estaría aprobando a estudiantes que no llegan al mínimo.

Si nos centramos en los alumnos, hay algún indicio que parece inducir a pensar que el rendimiento tiende a ser, en general, relativamente constante y sin discontinuidades. En un estudio realizado por Broc, Gil y Tejedor (2006), mediante un procedimiento de regresión logística binaria se pudo predecir en aproximadamente un 90% el número de suspensos de los alumnos en la tercera evaluación a partir de los obtenidos en la primera y un test de motivación hacia el estudio (Harter, 2012). No obstante, podría darse el caso que ciertos estudiantes hayan pasado por algún problema puntual y pasajero (enfermedad, problemas familiares, etc.), y que posteriormente se han recuperado, volviendo a su línea previa. También es posible que ciertos alumnos tengan “suerte” en las cuestiones que se les exigen en las pruebas, o que desarrollen habilidades de “copia” que son informadas de forma creciente por bastantes alumnos, lo que desembocaría en una evaluación falsa que no reflejaría la realidad de los conocimientos y competencias adquiridas por aquellos.

En principio parece haber suficiente evidencia para pensar que existen métodos estadísticos que nos permiten predecir con suficiente nivel de confianza el promedio de suspensos o aprobados que los alumnos/as obtendrían según los pronósticos, en relación con los obtenidos realmente. El reto de futuras investigaciones sería poder diagnosticar no en qué sujetos, que eso si puede hacerse mediante esta metodología del análisis de regresión, sino en qué asignaturas sería más probable que esa diferencia entre los valores reales y pronosticados pudiera darse. No sabemos si existen procedimientos estadísticos en esta dirección, pero sí podría afirmarse que, de forma intuitiva, esas discordancias habría que buscarlas, sobre todo, en los estudiantes cuyas calificaciones finales en las asignaturas correspondientes fueran iguales a cinco (cuyos pronósticos indican una puntuación igual o menor que cuatro). Asimismo también sería conveniente analizar las puntuaciones de los alumnos mejor dotados académicamente, cuyas calificaciones finales son nueves y dieces, explorando si las correspondientes al nueve han sido infladas hacia el diez, ya que este fenómeno, de hecho, sucede en la realidad. Estas prácticas

apelan a la honestidad y sensatez del profesorado y tienen, de hecho, implicaciones posteriores, de cara a la selectividad, en la obtención de unas medias superiores, que pueden perjudicar a otros alumnos que pertenecen a otros centros donde esas “inflaciones” no se han dado o se han dado en menor proporción.

También creemos que las notas cuantitativas del bachillerato deberían de incluir un decimal porque las evaluaciones serían más justas y precisas y se evitaría, al menos en parte, esas decisiones de subir al entero superior o bajar al entero inferior, que provocan desajustes en los pronósticos y decisiones del profesorado, a veces equivocadas.

Llegados a este punto sería deseable relacionar estas cuestiones con los principios éticos y deontológicos de la práctica docente, o de lo que actualmente se denomina “*buenas prácticas docentes*”. ¿Es justo que en unos Centros algunos profesores evalúen por encima de lo que les corresponde a sus alumnos por cuestiones personales, afectivas u organizativas? ¿Esto podría perjudicar a otros estudiantes de otros centros que con mejor desempeño reciben peores calificaciones? Claramente esta cuestión se va acumulando hasta llegar al final de segundo de bachillerato y, posteriormente, con el acceso a estudios de grado en los que las décimas y centésimas determinan de forma definitiva la entrada o no en las carreras en las que se han puesto muchas ilusiones por parte de los alumnos y sus familias. No obstante, hasta que esta ética y deontología profesional no se generalice y se interiorice en todo el profesorado, este tema tendrá una complicada solución.

References:

- Álvaro, M., Bueno, M. J., Calleja, J. A., Cerdán, J., Echevarría, M. J., García, C., Gaviría, J. L., Gómez, C., Jiménez, S., López, B., Martín-Jabato, L., Mínguez, A. L., Sánchez, A. y Trillo, C. (1990). *Hacia un modelo causal del rendimiento académico*. Madrid: Centro de Publicaciones del Ministerio de Educación y Ciencia: CIDE. Colección: Investigación, nº 43.
- Borkowski, J. G., y Thorpe, P. K. (1994). Self-Regulation and Motivation: A life-span perspective on Underachievement. En D. H. Schunk y B. J. Zimmerman (Eds.). *Self-Regulation of Learning and Performance. Issues and Educational Applications*, (pp. 45-74). Lawrence Erlbaum Associates, Publishers. Hillsdale. New Hersey. Hove, UK.
- Broc, M. Á., Gil, C., y Tejedor, J. A. (2006). La regresión logística binaria en la predicción del rendimiento académico en alumnos de bachillerato LOGSE. En *ANNALES (Anuario del Centro de la Universidad Nacional de Educación a Distancia, de Barbastro)*. Vol. XVIII (p. 179-218).
- Broc, M. Á. (2008). Predicción del rendimiento académico en alumnos de ESO y Bachillerato mediante el Inventario Clínico para adolescentes de Millon (escala MACI). *Anales de Psicología*, 24, (1), 158-167.

- Broc, M. Á. (2012). Influencia relativa de variables metacognitivas y volitivas en el rendimiento académico de estudiantes de Bachillerato (LOE). *Revista Española de Orientación y Psicopedagogía (REOP)*, 23, 3, 3º cuatrimestre, 63-80.
- Broc, M. Á. (2015). A Longitudinal Study of Academic Success and Failure in Compulsory Secondary Education and Baccalaureate Students through the Millon Adolescents Clinical Inventory (MACI). *Psychology*, 6, 1427-1437. Recuperado de <http://www.scirp.org/journal/psych>
<http://dx.doi.org/104236/psych.2015.612139>
- Gaviria, J. L. (2005). La equiparación del expediente de Bachillerato en el proceso de selección de alumnos para el acceso a la universidad. *Revista de Educación*, 337, 351-387.
- Gaviria, J. L., y Castro, M. (2005b). *Modelos Jerárquicos Lineales*. Madrid: La Muralla.
- González-Vallinas, P., Álvarez, J. L., Peiró i Gregori, S., y San Fabián Maroto, J. L. (2010). Mito y realidad en las diferencias de resultados en Bachillerato y en la prueba de acceso a la universidad de los centros públicos y privados en Asturias en el período 2004-2008. *Revista Iberoamericana de Educación*, 53 (7).
- Governa, M. A., López, M. A., y Pastor, J. T. (1987). La predicción del rendimiento como criterio para el ingreso en la universidad. *Revista de Educación*, 283, 235-248.
- Harter, S. (2012). *The Construction of the Self. Developmental and Sociocultural Foundations* (second edition). New York: The Guilford Press. <http://www.guilford.com>
- Ley Orgánica 8/2013, de 9 de diciembre, para la mejora de la calidad educativa. *Boletín Oficial del Estado*, n1 295, de martes 10 de diciembre de 2013.
- Martín, E., Martínez-Arias, R., Marchesi, A., Pérez, E. (2008). Variables that Predict Academic Achievement in the Spanish Compulsory secondary Education System: A Longitudinal Multi-Level Analysis. *The Spanish Journal o Psychology*, 11, (2), 400-413.
- Martínez-Arias, R. (2009). Usos, aplicaciones y problemas de los modelos de valor añadido en educación. *Revista de Educación*, 348, 217-250.
- Martínez-Arias, R., Gaviria, J. L., y Castro, M. (2009). Concepto y evolución de los modelos de valor añadido en educación. *Revista de Educación*, 348, 15-45.
- MedCalc (Version 16.4.3). (1993-2016). *MedCalc Software bvba*. Microsoft Partner. Recuperado de <http://www.medcalc.org>
- Miñano, P. (2009). *Un modelo causal-explicativo sobre la incidencia de las variables cognitivo-motivacionales en el rendimiento académico*. (Tesis

Doctoral). Departamento de Psicología Evolutiva y Didáctica. Facultad de Educación. Universidad de Alicante.

Moreno-Herrero, D., Sánchez-Campillo, J., y Jiménez-Aguilera, J. (2014). ¿En los centros privados se inflan las calificaciones de los estudiantes? *Revista de Educación*, 366, Octubre-diciembre, 267-289. DOI: 10.4438/988-592X-RE-2014-366-275

Orden de 14 de octubre de 2008, del Departamento de Educación, Cultura y Deporte, sobre la evaluación en Bachillerato en los centros docentes de la Comunidad Autónoma de Aragón. *Boletín Oficial de Aragón*, nº 77, de 28 de octubre de 2008.

Pavía, J. M. y Vila, E. (2010). ¿Equidad académica en el acceso a la universidad? ¿Enseñanza secundaria pública versus privada? *Investigaciones de Economía de la educación*, 4, 87-94.

Real Decreto 1467/2007, de 2 de noviembre por el que se establece la estructura del bachillerato y se fijan sus enseñanzas mínimas. *Boletín Oficial del Estado* núm. 266, del Martes 6 de noviembre de 2007.

Ruiz de Miguel, C. (2009). Las escuelas eficaces: un estudio multinivel de factores explicativos del rendimiento escolar en el área de matemáticas. *Revista de Educación*, 348, 355-376.

Santrock, J. W. (2011). *Psicología de la Educación* (4ª edición). Madrid: McGrawHill.