

Los patrones atípicos de respuesta una fuente de información para evaluar la calidad del proceso enseñanza-aprendizaje

Eduardo Doval¹, M. Dolors Riba¹, Marta Fuentes², Jordi Renom³,

¹ Departament de Psicobiologia i de Metodologia de les Ciències de la Salut (UAB),

² Departament de Psicologia Bàsica, Evolutiva i de l'Educació (UAB), ³ Departament de Psicologia Social i Psicologia Quantitativa (UB)

1. INTRODUCCIÓN

- Los exámenes tipo test constituyen un recurso habitual en la evaluación de los aprendizajes.
- La puntuación obtenida en estas pruebas suele utilizarse como indicador de la adquisición de los aprendizajes.
- La validez de este indicador se basa en el supuesto de que el alumno contestará correctamente las preguntas que evalúen un nivel de conocimientos inferior o igual al suyo, e incorrectamente las que evalúen niveles de conocimientos superiores.
- El nivel de dificultad de las preguntas (proporción de respuestas correctas en el grupo de referencia) se toma como un indicador del nivel de conocimiento evaluado.
- Se espera que si un alumno comete errores, éstos se acumulen en las preguntas más difíciles. Los estudiantes que se alejan de esta pauta contestan la prueba con un patrón de respuesta atípico (PAR).
- La presencia de PAR puede identificarse con el Índice de Precaución Modificado (IPM; Harnish y Linn, 1985).

ÍNDICE DE PRECAUCIÓN MODIFICADO (IPM)

- Es uno de los indicadores no paramétricos que mejor detecta la presencia de PAR (Karabatsos, 2003).
- Se basa en la covariación del patrón de respuestas, (1=respuesta correcta, 0=respuesta incorrecta) con el vector de dificultades de los ítems.
- Toma valores entre 0 y 1. Los valores altos (e.g. superiores a 0.50) indican presencia de PAR.
- Puede calcularse con el paquete *PerFit* de R (Tendeiro, 2016).

2. OBJETIVO

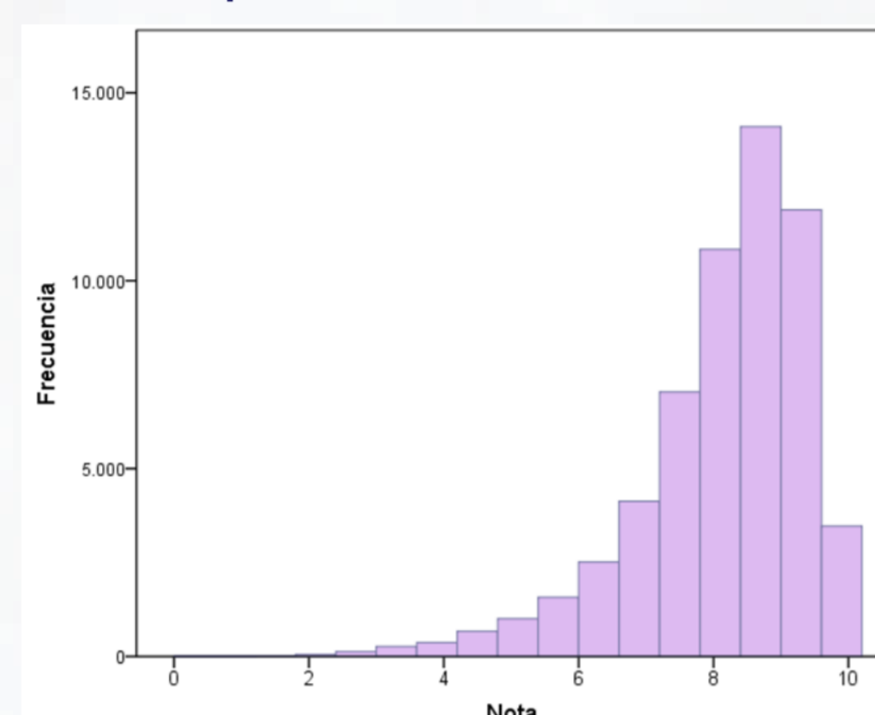
Identificar la presencia y el tipo de PAR en una prueba de evaluación de aprendizajes, con el objetivo final de proporcionar evidencia diagnóstica del proceso evaluación-aprendizaje.

3. MÉTODO

Se analizan, con una metodología propia, las respuestas a la prueba de competencias básicas de castellano (33 preguntas con 3 opciones de respuesta) realizada por el *Consell Superior d'Avaluació del Sistema Educatiu de Catalunya* durante el curso 2014-15 a todos los alumnos de cuarto curso de la E.S.O.

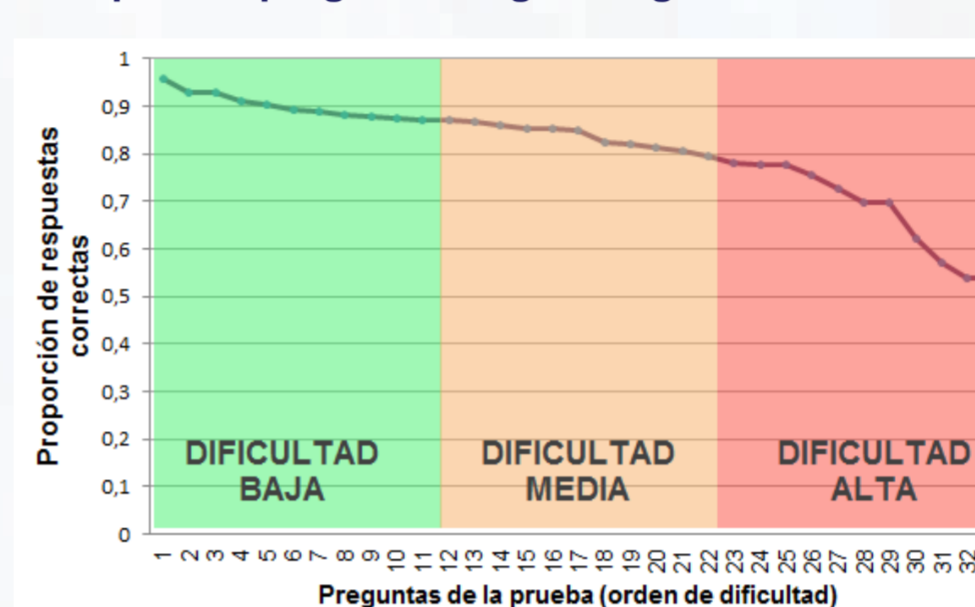
Respondieron 58.072 alumnos. La figura 1 muestra la distribución de las puntuaciones (sobre 10 puntos).

Figura 1
Distribución de puntuaciones de los alumnos evaluados



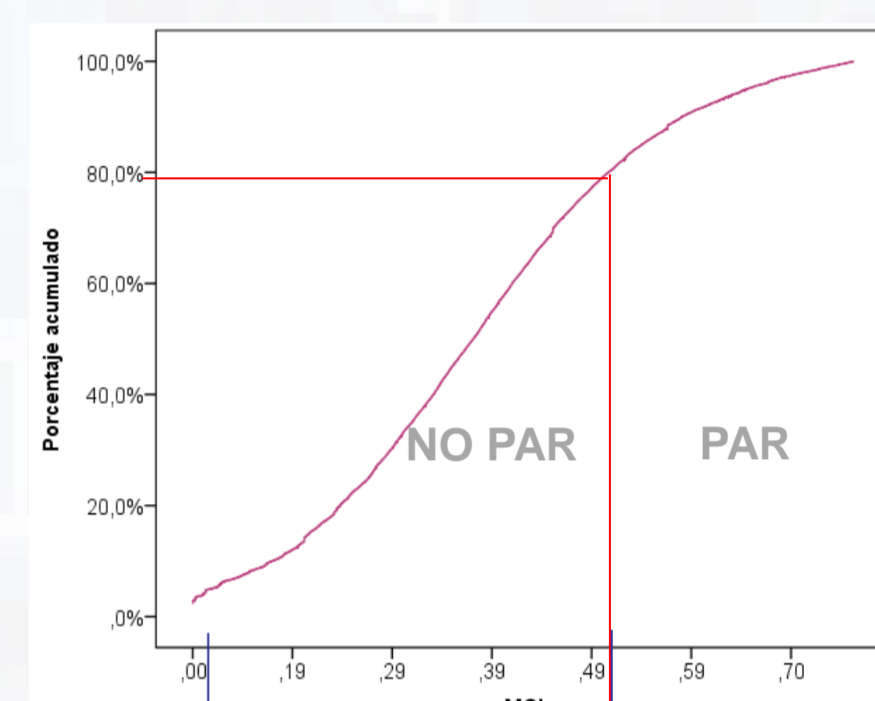
- Se calcula el grado de dificultad de las preguntas de la prueba a partir de la proporción de respuestas correctas en el grupo de referencia. Las preguntas se ordenan de más fáciles a más difíciles (ver perfil de dificultad en la figura 2)
- En función de los centiles (33 y 66) se identifican tres bloques de preguntas según su nivel de dificultad -baja, media y alta- (figura 2).

Figura 2
Bloques de preguntas según su grado de dificultad



- Se calcula el IPM para todas las pautas de respuesta. Las pautas con IPM > 0.5 -el 21% de las pautas más extremas- se identifican como PAR (figura 3).

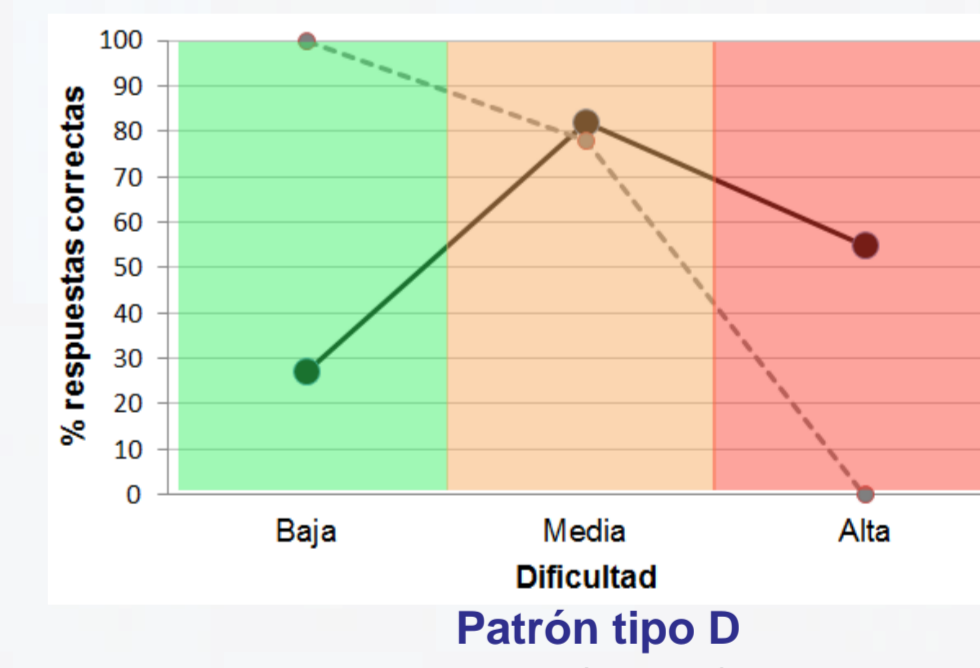
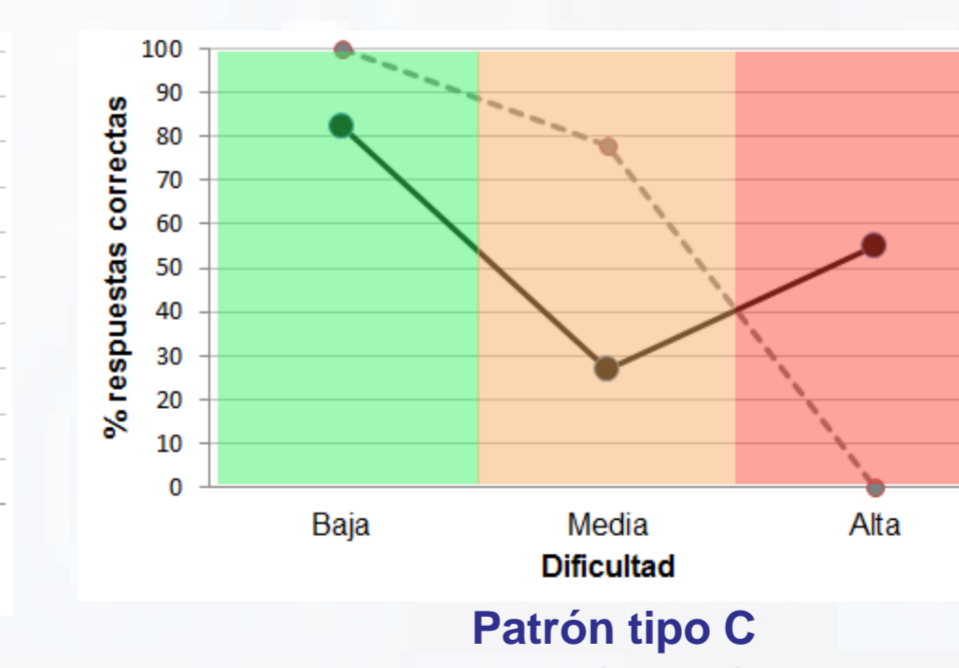
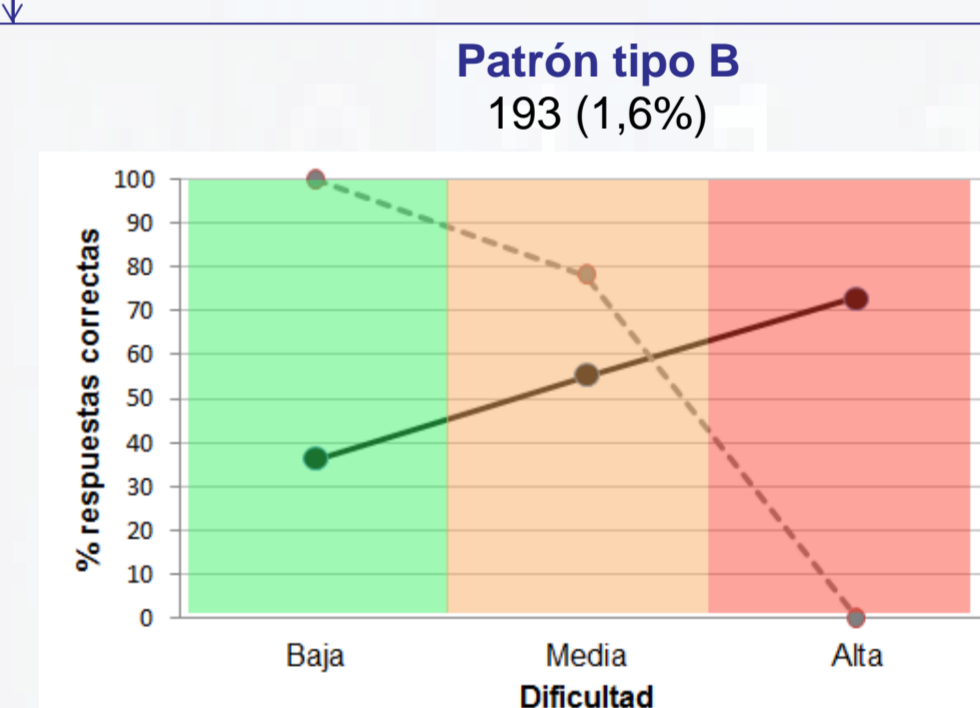
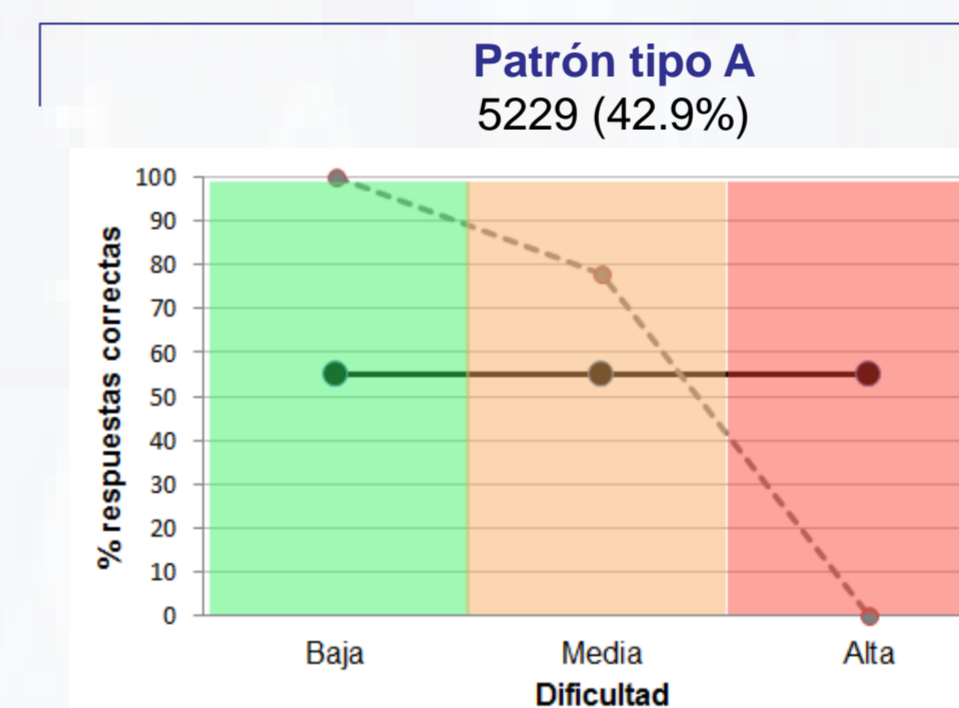
Figura 3
Distribución acumulada de valores del IPM



- Se identifican los tipos o patrones de PAR según la tendencia del porcentaje de respuestas correctas en cada bloque de preguntas (línea continua) y se comparan con el patrón esperado por puntuación (línea discontinua).

4. RESULTADOS

A modo de ejemplo, se ilustran los tipos de patrones de respuesta identificados con cinco alumnos que comparten puntuación (5.5). En línea punteada se muestra el patrón esperado para esa nota.



5. CONCLUSIONES

De entre los alumnos seleccionados por haber contestado con PAR (el 21% del total), el 42.9% contestó con un patrón de respuestas similar en cada bloque (Patrón A). El resto contestó correctamente más preguntas de las esperadas en el bloque de respuestas más difíciles, compensando este desvío con menos respuestas correctas de lo esperado en el bloque con preguntas de dificultad moderada (Patrón C, 24,5%), baja (Patrón D, 24%) o en ambos (Patrón B, 1,6%). En este último caso el patrón de respuestas es completamente contrario al esperado.

Una misma puntuación puede proceder de patrones de respuesta muy diferentes. Por este motivo, las inferencias sobre el nivel de aprendizaje no deben realizarse únicamente a partir del resultado final de la prueba. Los PAR identificados pueden resultar de conductas ilícitas (e.g. producto de la copia de respuestas), de malas pautas de estudio (e.g. centrar el estudio en los conceptos más difíciles y desatender los más sencillos) o de estrategias al realizar la prueba (e.g., concentrarse únicamente en las preguntas más complejas). En los dos últimos casos el potencial de los alumnos evaluados pueden ser mayor de lo que refleje su nota final. Es responsabilidad de los docentes utilizar la información proporcionada por el tipo de PAR para a precisar el nivel del alumno evaluado y de esa manera mejorar la evaluación del proceso de enseñanza-aprendizaje.

REFERENCIAS

- Harnisch, D. L. y Linn, R. L. (1981) Analysis of item response patterns: Questionable test data and dissimilar curriculum practices. *Journal of Educational Measurement*, 18(3), 133-146.
- Karabatsos, G. (2003) Comparing the Aberrant Response Detection Performance of Thirty-Six Person-Fit Statistics. *Applied Measurement in Education*, 16(4), 277-298.
- Tendeiro, J. (2016). Package "PersonFit", v. 1.4.1. <http://cran.r-project.org/web/packages/PerFit/PerFit.pdf>

Este trabajo ha recibido el apoyo económico de la Dirección General de Investigación y del Plan Nacional de I+D+i, del Ministerio de Economía y Competitividad, con el proyecto EDU2013-41399-P.

Agradecemos al *Consell Superior d'Avaluació de la Generalitat de Catalunya* habernos facilitado los datos de este estudio.