

Detección de la presencia de patrones atípicos de respuesta asociados a bajo rendimiento: propuestas para una evaluación formativa

Eduardo Doval¹ y M. Dolors Riba¹

¹Departament de Psicobiologia i de Metodologia de les Ciències de la Salut. Facultat de Psicologia, Universitat Autònoma de Barcelona.

eduardo.doval@uab.cat; dolors.riba@uab.cat

INTRODUCCIÓN

En las pruebas tipo test, la presencia de determinados **patrones atípicos de respuesta (PAR)** da lugar a puntuaciones que pueden conducir a inferencias incorrectas sobre el nivel real de conocimientos o competencias que posee el estudiante evaluado.

A nivel formativo, la presencia de PAR cobra especial importancia en aquellos casos en que el rendimiento mostrado en las pruebas infravalora su nivel real.

La identificación de estos patrones es necesaria para la detección de aquellos aspectos (motivacionales, de pautas de estudio o de otro tipo) que limitan el rendimiento.

OBJETIVOS

- 1: Describir la presencia de patrones atípicos de respuesta y evaluar su impacto sobre la evaluación de competencias
- 2: Analizar si se observa una constancia individual en más de una prueba de evaluación

MÉTODO DE INVESTIGACIÓN

Se analizan respuestas a dos pruebas realizadas por el Consell Superior d'Avaluació de Catalunya para comprobar las competencias lingüísticas básicas en lengua catalana y castellana.

Población

Alumnado de 4º de ESO de Cataluña matriculado en el curso 2013-14.
64.505 alumnos de colegios públicos, concertados o privados.

56.557 alumnos (87.7%) realizaron la prueba de catalán.

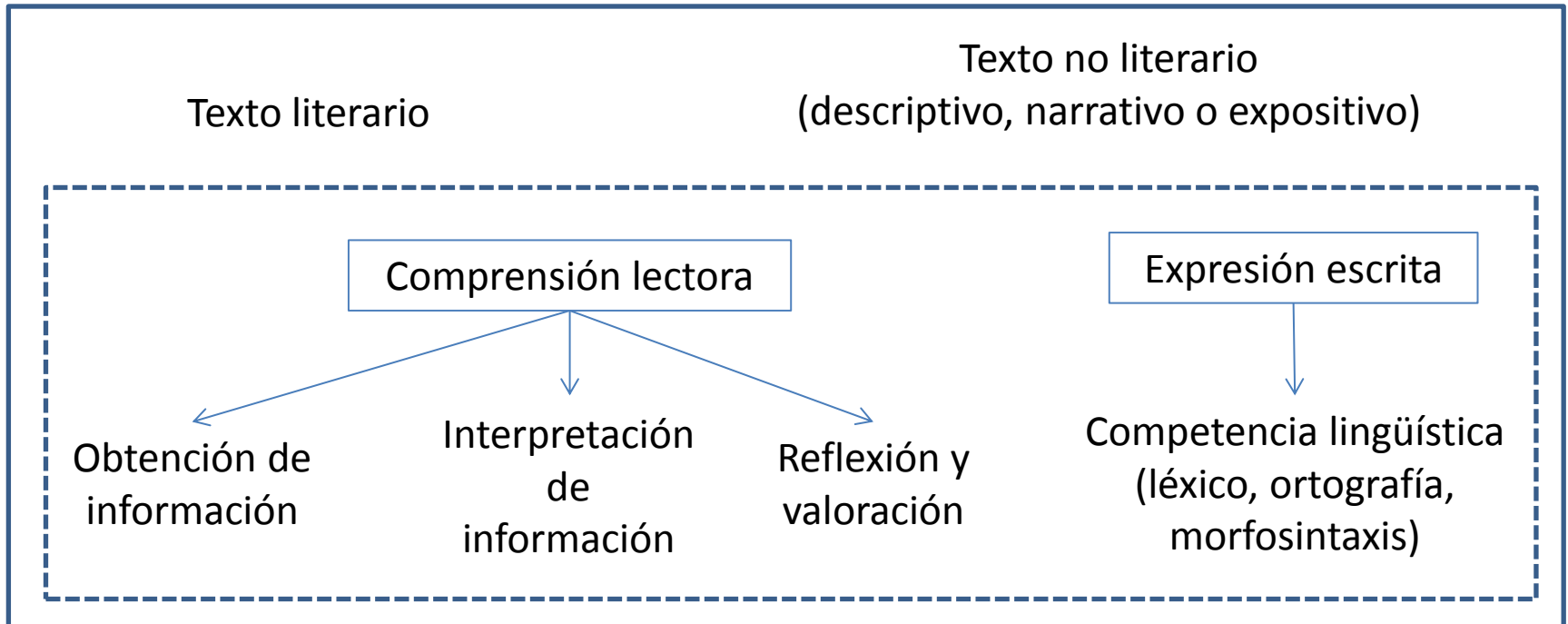
56.450 alumnos (87.5%) realizaron la prueba de castellano.

55.093 alumnos (85.4%) realizaron las dos pruebas.

MÉTODO DE INVESTIGACIÓN

Instrumentos

Pruebas de competencias básicas en lengua catalana y castellana.



Ambas pruebas tienen 32 preguntas tipo test (4 opciones o V/F).

MÉTODO DE INVESTIGACIÓN

Análisis

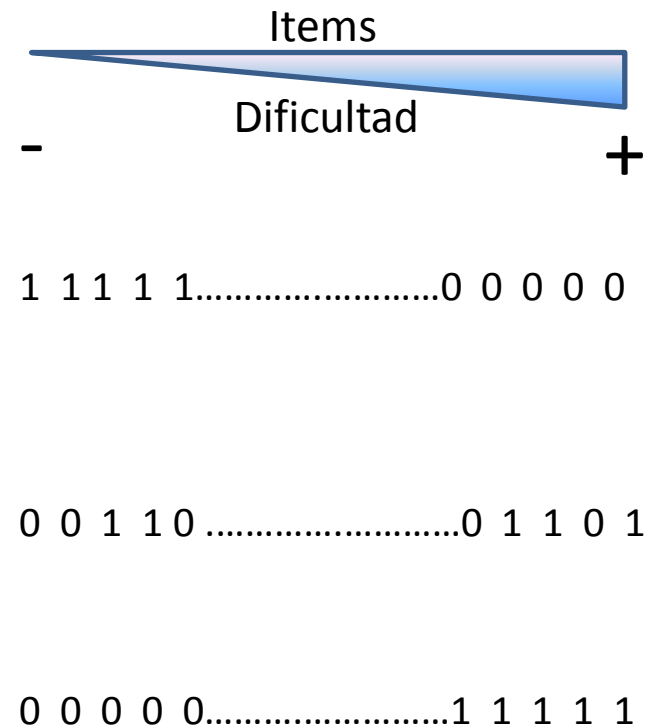
Los PAR se han identificado con el Índice de Precaución Modificado (MCI).

Para analizar las pautas de respuesta es necesario ordenar los ítems por su nivel de dificultad.

Una pauta perfecta correspondiente a una puntuación k , es aquella en que se solo se responden correctamente los k ítems más fáciles. A una pauta perfecta le corresponde un $MCI = 0$.

Un PAR es aquella forma de responder que se aleja de una pauta perfecta.

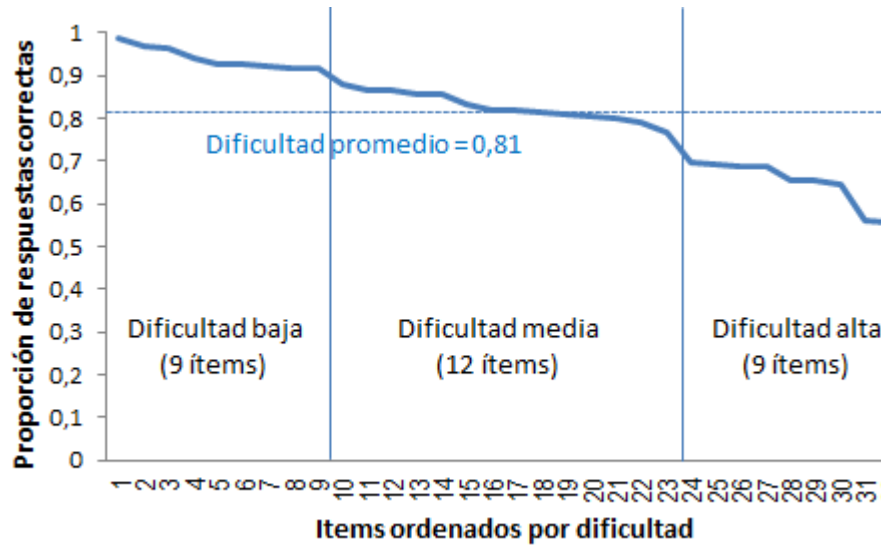
Una pauta totalmente atípica es aquella en que sólo se aciertan los k ítems más difíciles. En este caso el valor de $MCI = 1$.



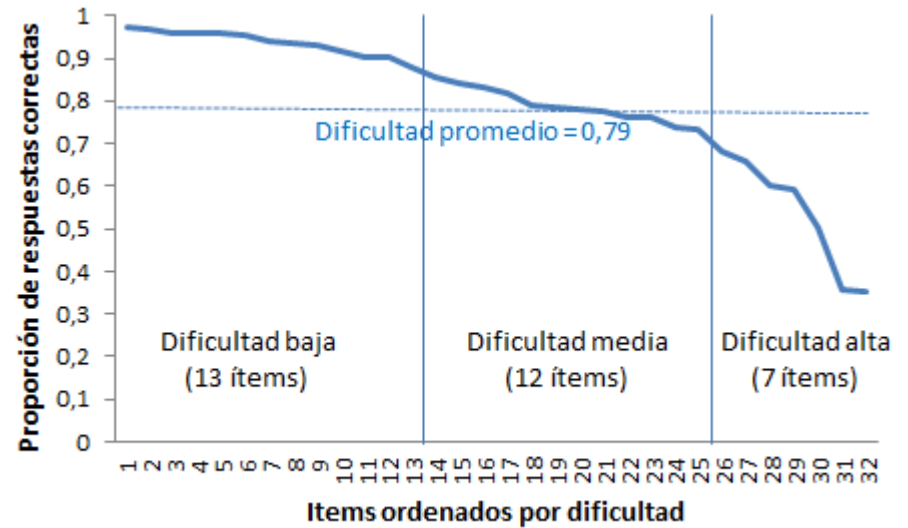
Los análisis se han realizado con el paquete de R PerFit (Tendeiro, 2015)

RESULTADOS

lengua catalana



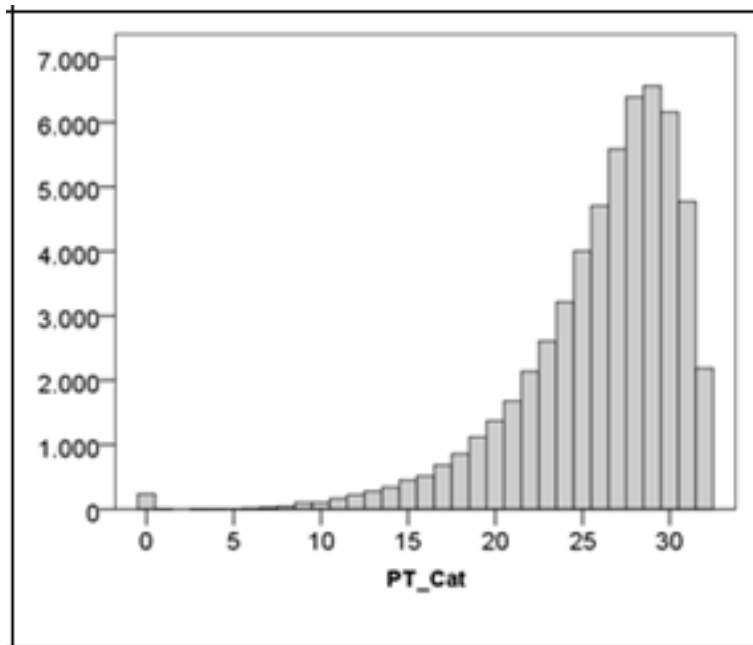
lengua castellana



RESULTADOS

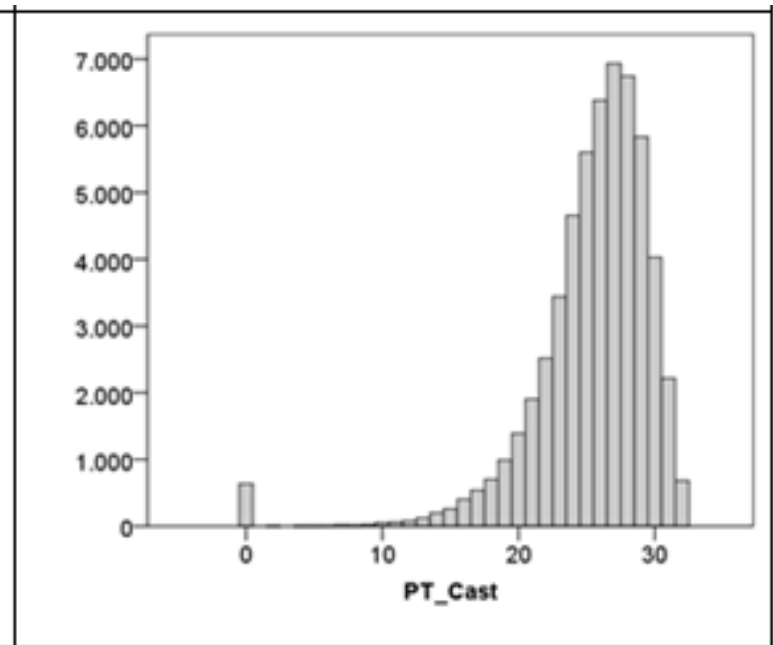
Puntuación total

lengua catalana



Media = 26,01
D.E.= 4,71

lengua castellana



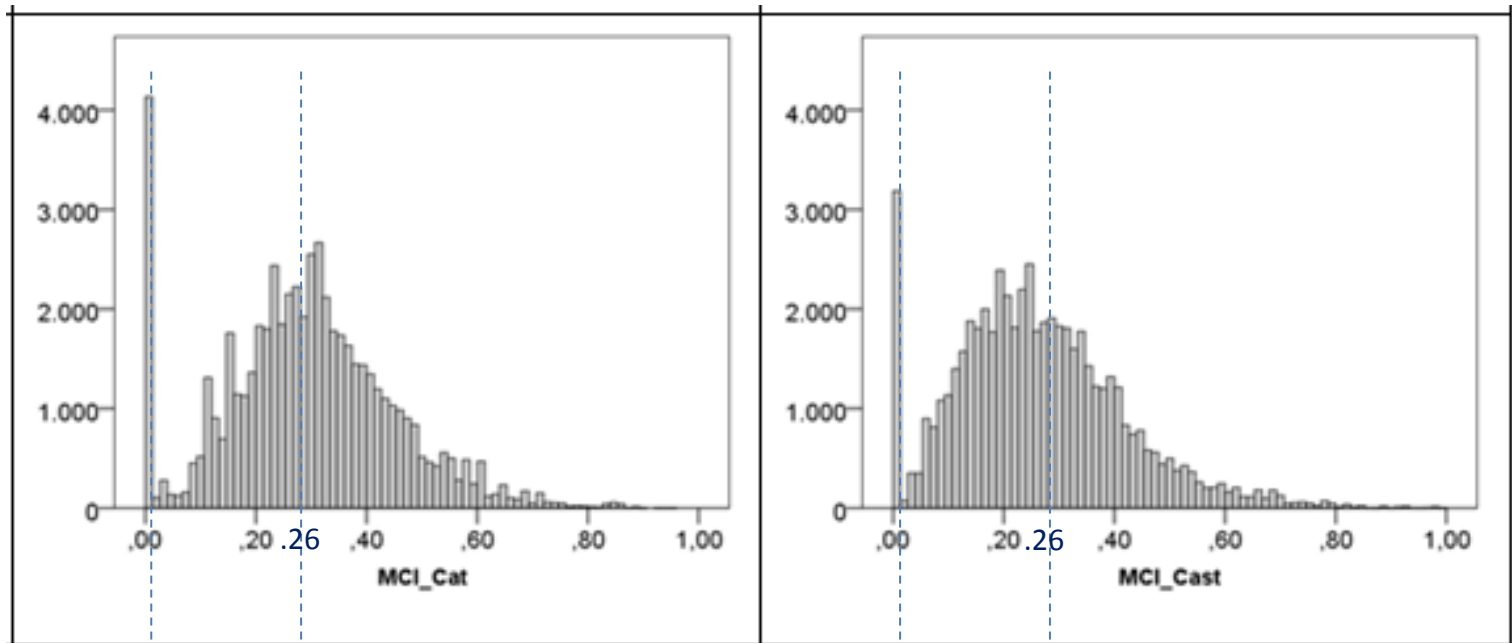
Media = 25,41
D.E.= 4,54

RESULTADOS

MCI

lengua catalana

lengua castellana



Pauta perfecta
(4,8%)

No PAR
(25,1%)

PAR
(59,2%)

Pauta perfecta
(4,5%)

No PAR
(49%)

PAR
(46,5%)





(1%)

(17,7%)

(27,9%)

RESULTADOS

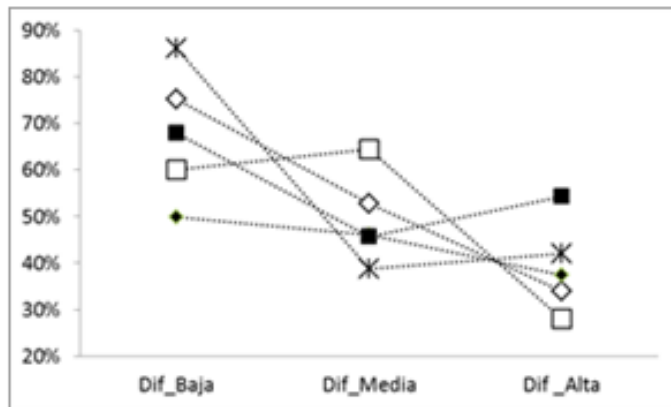
Segmentación

| | | Nota 4 - 6 | | MCI > 0,26 |
|-------------------|---|-------------------|---|----------------------|
| Catalán |  | 4242 |  | 2340 (55,2%) |
| Castellano |  | 3220 |  | 1447 (44,9%) |

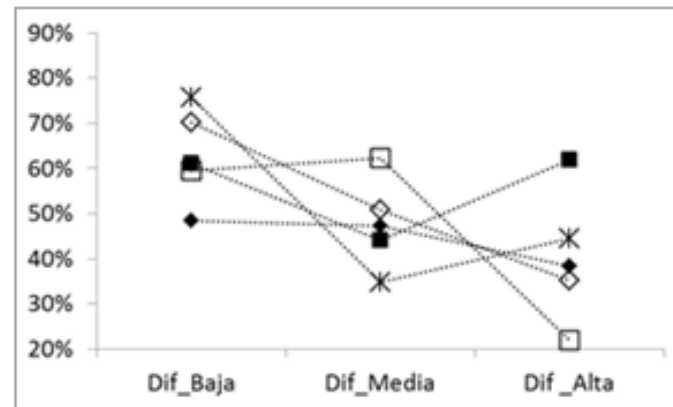
RESULTADOS

Análisis de clústers

Catalán



Castellano



.....◆..... Patrón 1 □..... Patrón 2 ◆..... Patrón 3 ■..... Patrón 4 ✕..... Patrón 5

| | | | | | |
|------------|----------------|----------------|----------------|----------------|----------------|
| Catalán | 955 (40,8%) | 298 (12,7%) | 231 (9,9%) | 347 (14,8%) | 509 (21,8%) |
| Castellano | 516 (35,7%) | 197 (13,7%) | 199 (13,8%) | 200 (13,8%) | 335 (23,2%) |

DISCUSIÓN

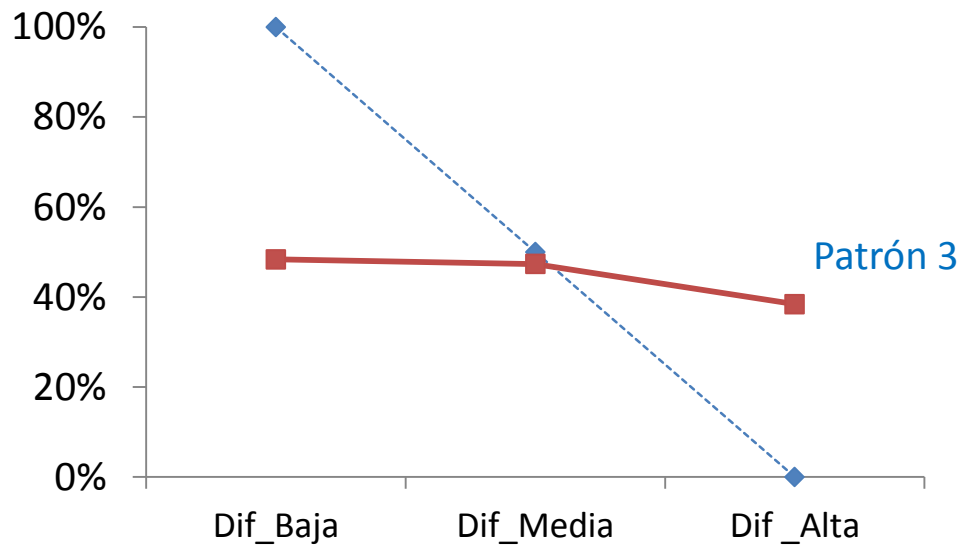
Nuestro principal objetivo era analizar la presencia de patrones atípicos de respuesta en pruebas tipo test en tanto que indicadores de procesos de respuesta inadecuados.

- Hemos comprobado que los patrones atípicos aparecen con una frecuencia considerable tanto en las pruebas de castellano (46,5%) como de catalán (59,2%).
- Además cabe destacar que un 27,9% de la población estudiada es consistente en la presentación de esta forma atípica de responder a las pruebas.

Hemos identificado cinco tipologías de respuesta atípicas.

DISCUSIÓN

Patrón 3

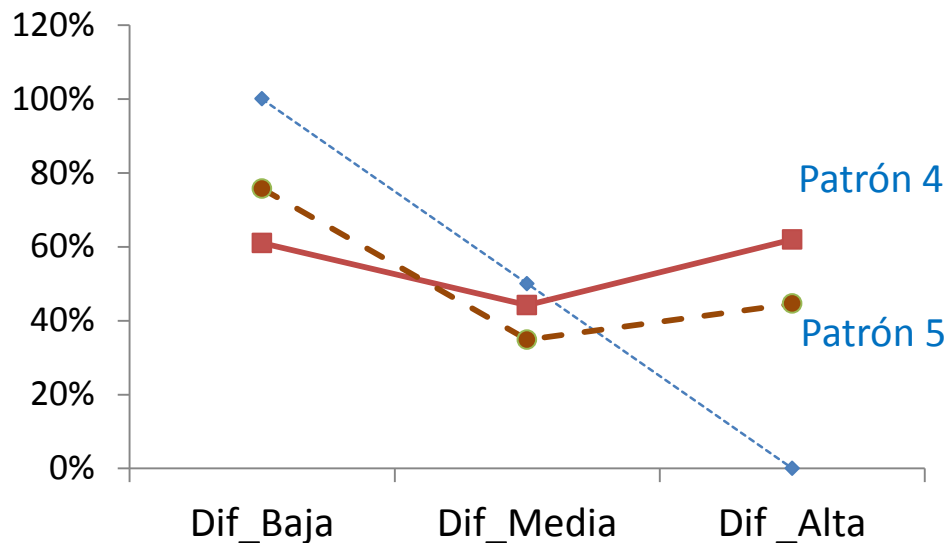


Porcentaje de respuestas correctas similar en los distintos tipos de ítems, independientemente de su dificultad.

Se corresponde con el que podría esperarse bajo el supuesto de que se responde a la prueba de forma aleatoria.

De ser cierto, este patrón no tendría mayor consecuencia sobre la aparente sobreestimación de la competencia si se aplicase alguna corrección a la puntuación en función del número de alternativas de respuesta de los ítems.

DISCUSIÓN



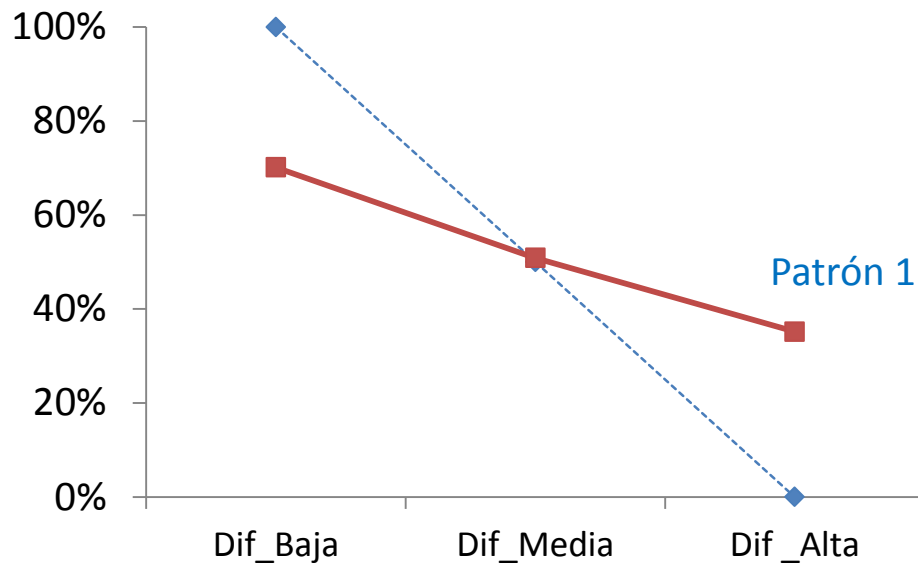
Patrones 4 y 5

El porcentaje de respuestas correctas a ítems de alta dificultad no es esperable dada la puntuación alcanzada en la prueba.

Estos perfiles son propios de personas que copian las respuestas de las preguntas difíciles. En este caso conducirían a una sobreestimación de la competencia evaluada

También pueden deberse a malos hábitos de estudio (desatender el estudio de conceptos más básicos de la materia). En estos casos, la estimación de la competencia estaría subestimada.

DISCUSIÓN

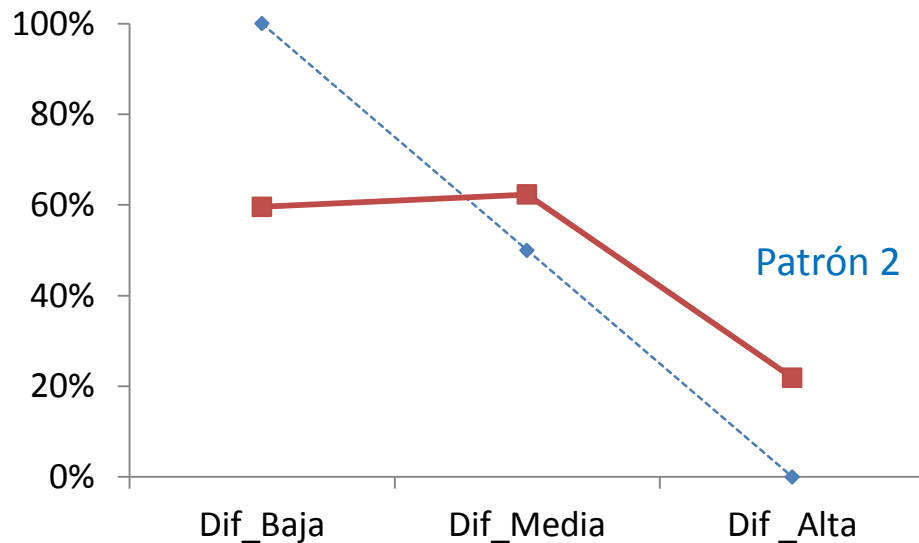


Patrón 1

Muestra una tendencia de respuestas correctas que decrece con la dificultad de los ítems.

Aunque se esperaría un mayor porcentaje de respuestas correctas en ítems fáciles y menos en ítems difíciles, no creemos que el desvío de la pauta perfectamente correcta conduzca, en este caso, a un sesgo en la puntuación final porque la tendencia de respuesta es la esperada.

DISCUSIÓN



Patrón 2

Se responden correctamente menos preguntas fáciles de las que cabría esperar dada la alta ejecución mostrada en las preguntas de dificultad intermedia.

Puede deberse a un problema asociado al proceso de enseñanza-aprendizaje o también a estrategias incorrectas a la hora de afrontar preguntas que pueden ser interpretadas como demasiado evidentes.

Subestima la competencia evaluada.

DISCUSIÓN

Nuestro objetivo secundario, ha sido analizar la presencia de consistencia individual.

- Un porcentaje elevado de estudiantes, el 46,6%, repite la manera en que responde a ambas pruebas, bien sea con presencia de PAR o sin ella.
- Este resultado da una idea de que la presencia o ausencia de PAR es un hecho relevante que ha de ser incorporado en el proceso de evaluación formativo, y más si cabe cuando el 23,7% de los casos que presentan PAR en ambas pruebas, coinciden en el tipo de patrón.

DISCUSIÓN

La interpretación de las evaluaciones con exámenes tipo test, no pueden basarse en el simple recuento de respuestas correctas.

Como se ha visto, un estudio de cómo y dónde se distribuyen los aciertos y los errores proporciona información de utilidad que puede ser aprovechada para mejorar las evaluaciones formativas.

MUCHAS GRACIAS POR SU ATENCIÓN

eduardo.doval@uab.cat

Esta investigación se ha realizado gracias a la ayuda de la Dirección General de Investigación y Gestión del Plan Nacional de I+D+I, del Ministerio de Economía y Competitividad, a través del proyecto EDU2013-41399-P.

Agradecemos al Consell Superior d'Avaluació de Catalunya su colaboración en este proyecto.