



Las puntuaciones obtenidas en los tests de conocimientos: ¿son siempre indicadores válidos del aprendizaje?

Eduardo Doval¹, M. Dolors Riba¹, Jordi Renom², Marta Fuentes³

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

²Departament de Metodologia de les Ciències del Comportament.
Facultat de Psicologia.
(UB)

³Departament de Psicologia Bàsica, Evolutiva i de l'Educació.
Facultat de Ciències de l'Educació.
(UAB)

INTRODUCCIÓN

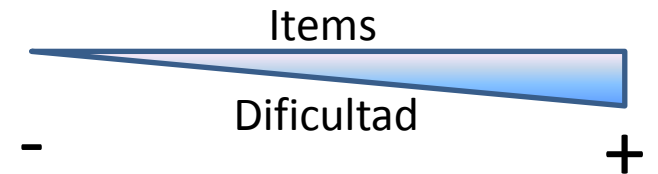
Las respuestas incoherentes a las preguntas de pruebas tipo test dan lugar a patrones de respuestas atípicos (PAR, Meijer y Sitjsma, 2001).

Estos patrones deberían ser identificados y estudiados individualmente con objeto de evaluar hasta qué punto proporcionan puntuaciones válidas.

También pueden proporcionar información útil sobre cómo ha sido el proceso de adquisición del aprendizaje.

INTRODUCCIÓN

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.

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.

1 1 1 1 1.....0 0 0 0 0
0 0 1 1 00 1 1 0 1
0 0 0 0 0.....1 1 1 1 1

OBJETIVOS

1: Cuantificar la presencia de Patrones Atípicos de Respuesta en la evaluación de competencias básicas de inglés en alumnos del último curso de la educación primaria.

2: Determinar las características del PAR y su asociación a diferentes niveles de rendimiento.

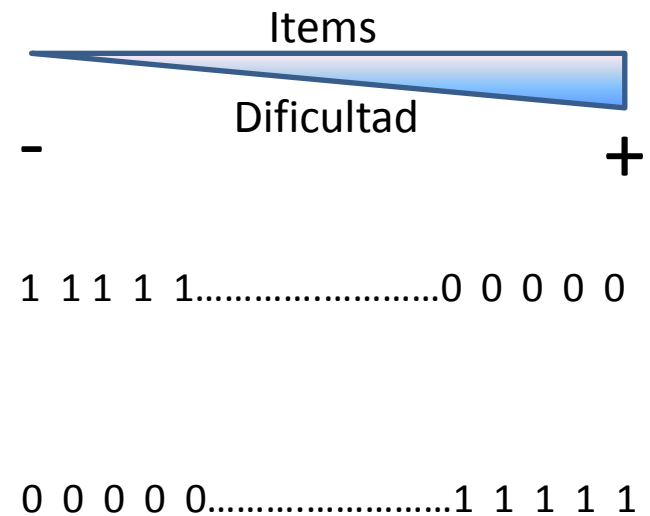
MÉTODO DE INVESTIGACIÓN

Análisis

Los PAR se han identificado con índice Iz^* (Snijders, 2001)

Valores positivos y altos de Iz^*

Valores de Iz^* inferiores a -1,64



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

MÉTODO DE INVESTIGACIÓN

Se analizan respuestas de una prueba realizada por el Consell Superior d'Avaluació de Catalunya para comprobar las competencias lingüísticas básicas en inglés de los alumnos de sexto curso de educación primaria.

Población

Alumnado de 6º de Educación Primaria de Cataluña matriculado en el curso 2013-14 (72.153 alumnos).

Realizaron la prueba de inglés 65.766 (91.1%) alumnos

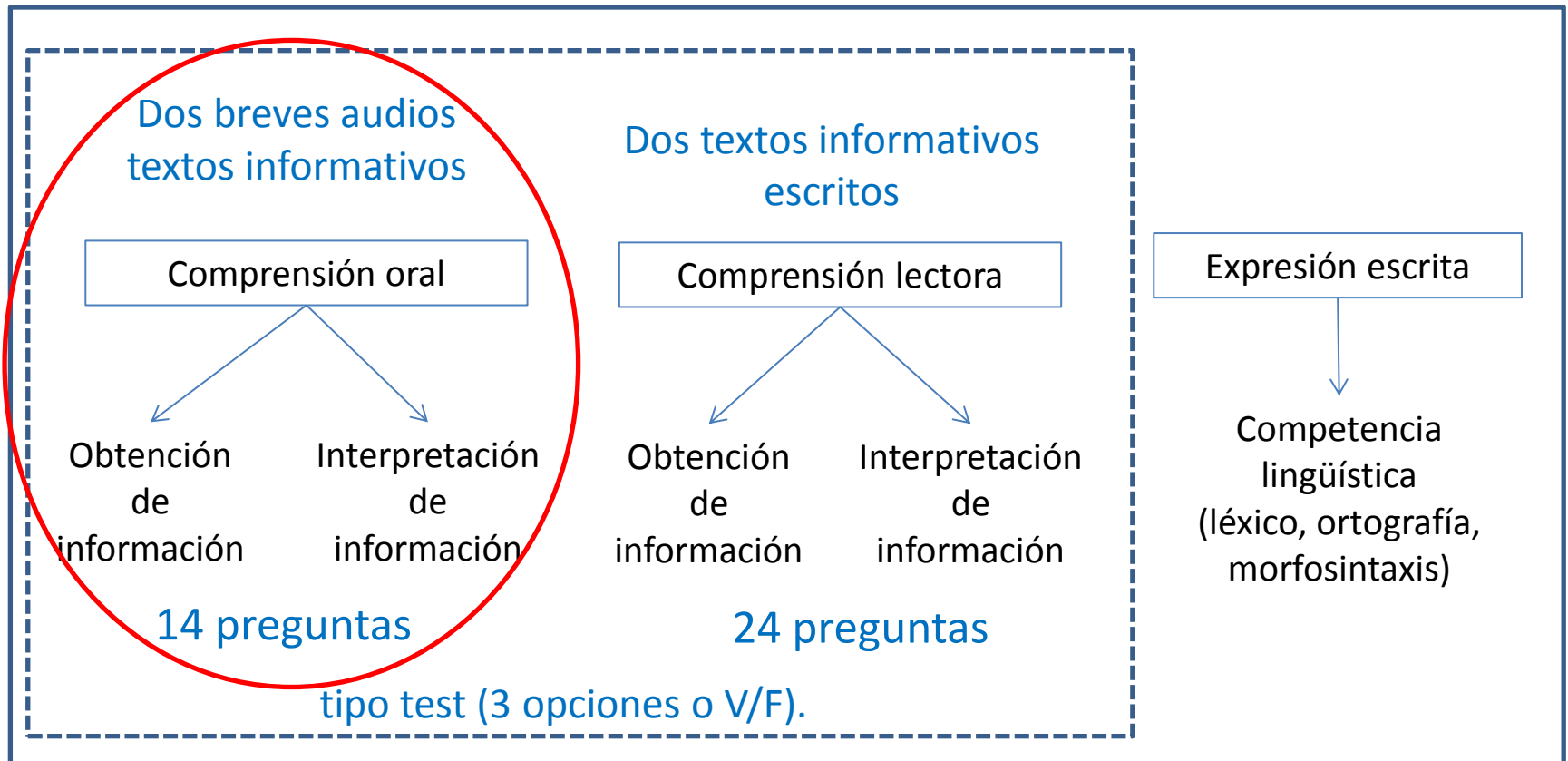
de 2182 centros

1585 públicos (73%)
572 concertados (26%)
25 privados (1%)

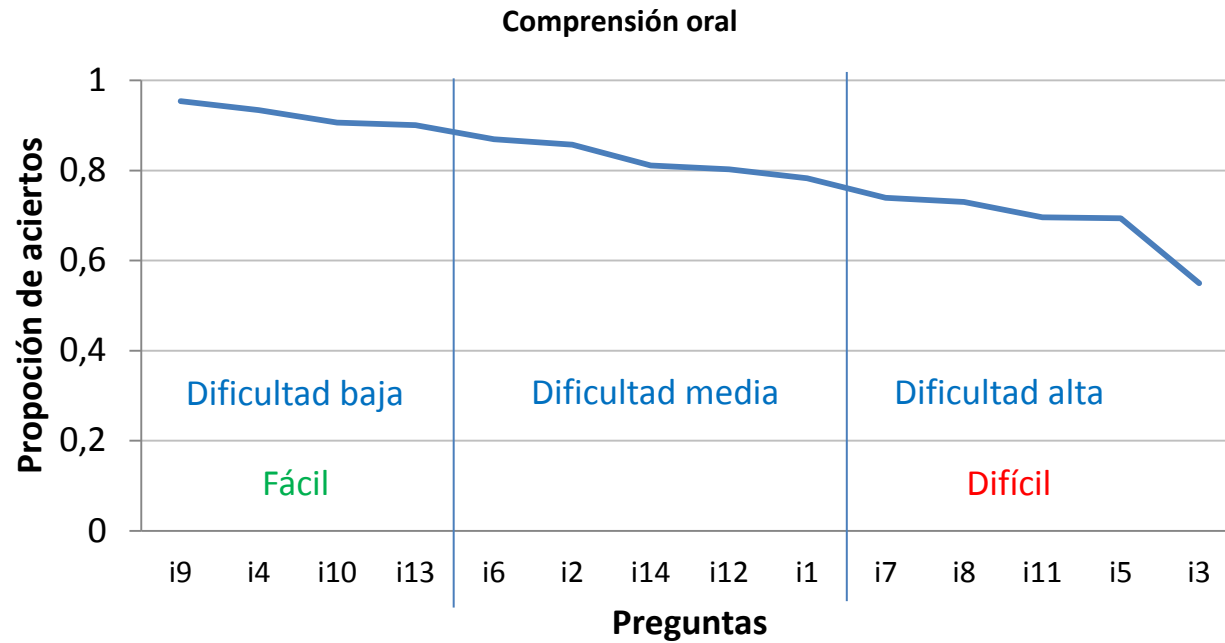
MÉTODO DE INVESTIGACIÓN

Instrumento

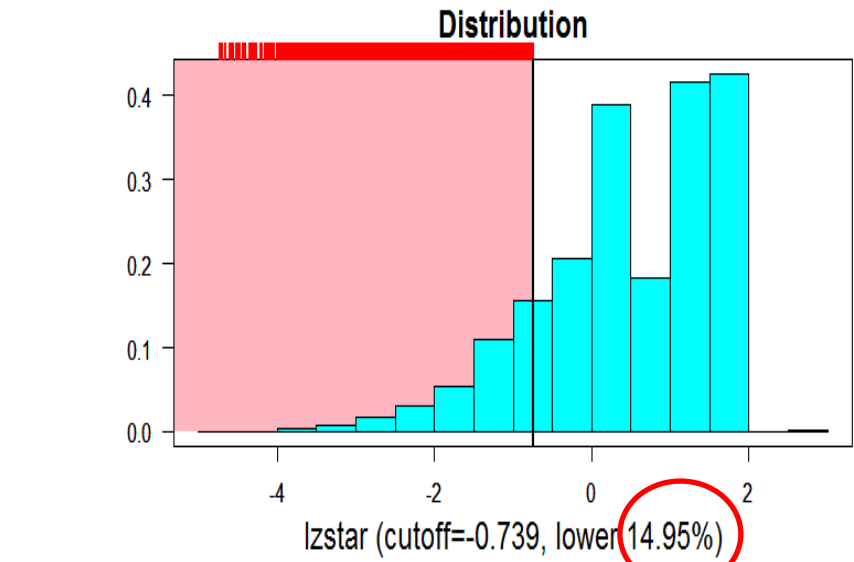
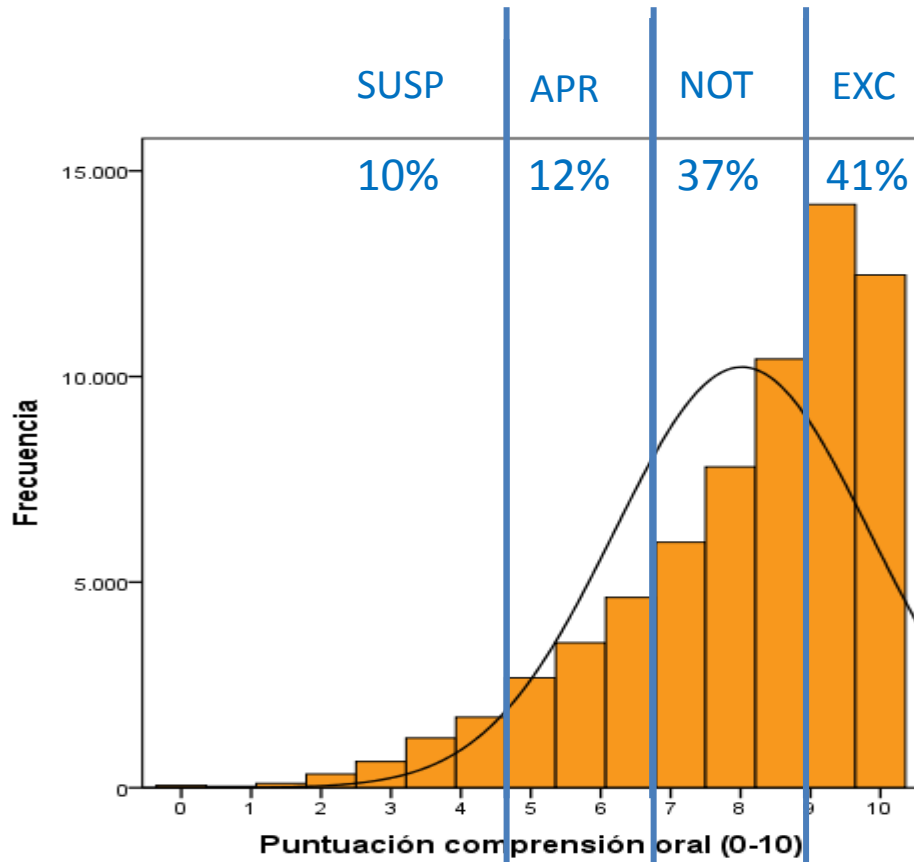
Prueba de competencias básicas en inglés.



RESULTADOS: Comprensión oral



RESULTADOS: Comprensión oral



Media = 8,02
Desviación estándar = 1,832

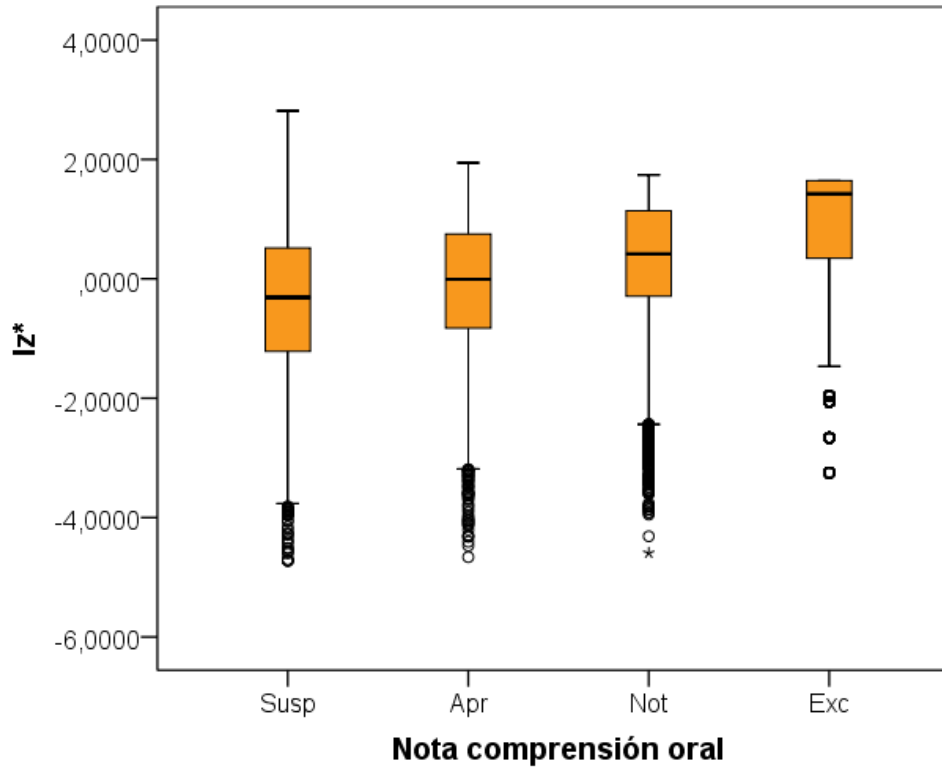
PAR → 25% 23% 38% 14%

37% 27% 16% 5%

El 14% de las PAR se concentran en las notas excelentes

El 5% de los excelentes tienen PAR

RESULTADOS: Comprensión oral

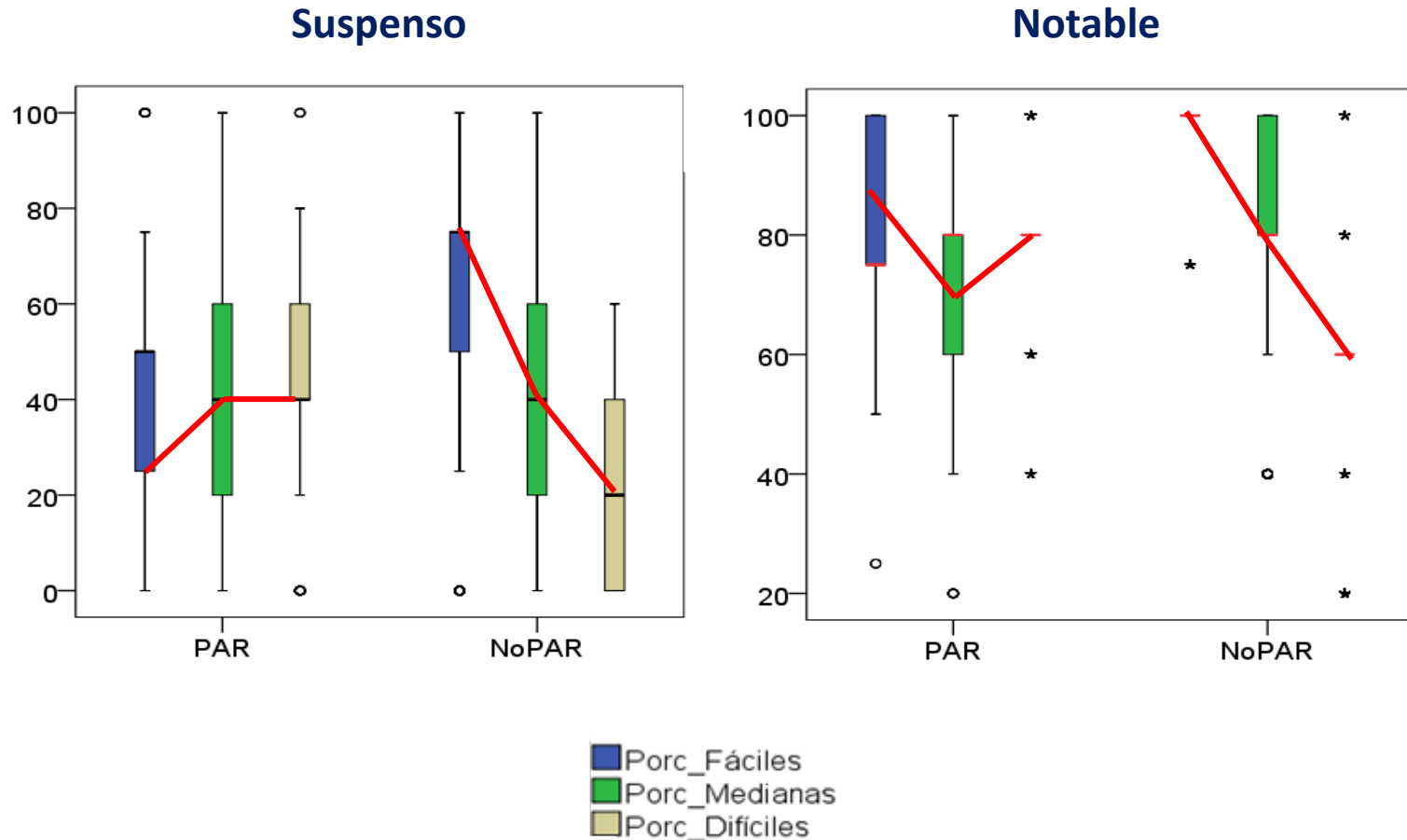


Correlación entre la puntuación total y el índice $Iz^* = 0.48$

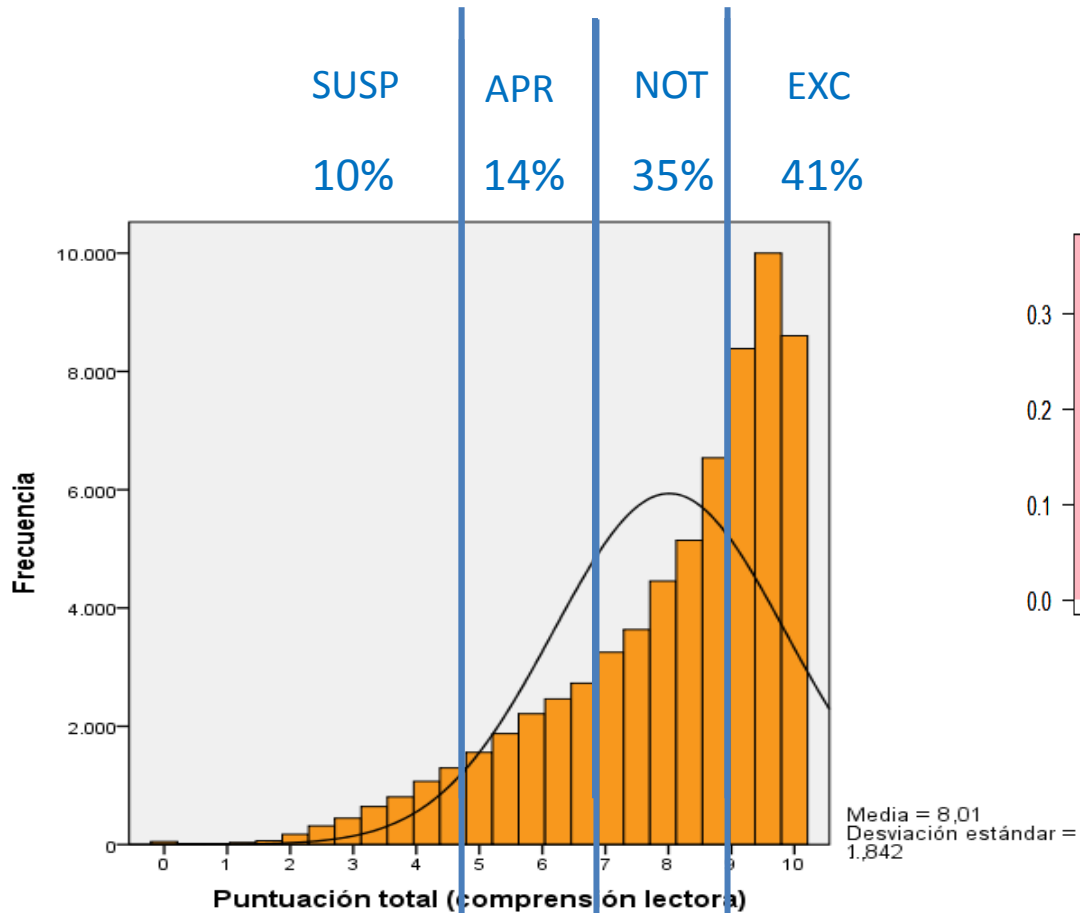
	Susp	Apr	Not	Exc
Mínimo	-4,72	-4,66	1,74	-3,25
Máximo	2,81	1,94	-4,59	1,64
Media	-0,39	-0,10	0,30	1,06
D.E.	1,26	1,11	0,97	0,87

RESULTADOS: Comprensión oral

Distribución de la proporción de aciertos a las preguntas de dificultad baja, moderada y alta en pautas con y sin PAR.

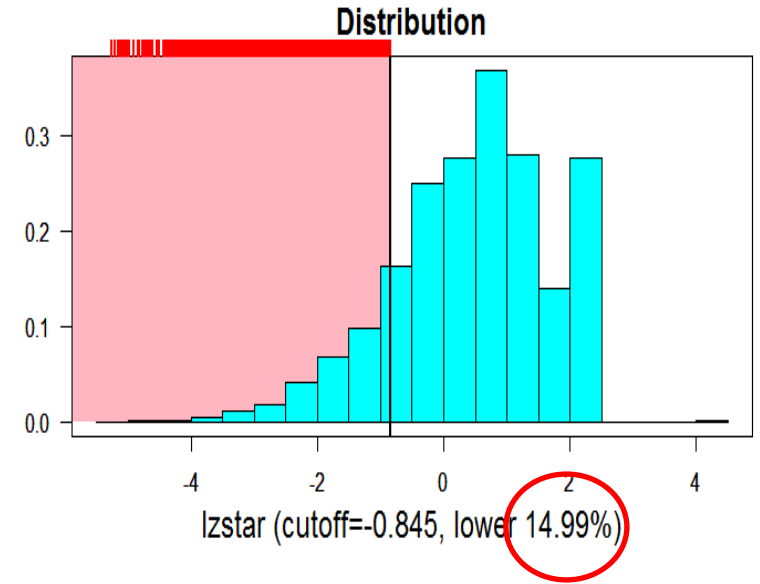


RESULTADOS: Comprensión lectora



SUSP 10%
APR 14%
NOT 35%
EXC 41%

PAR → 28%
42% 21% 13% 8%



RESULTADOS: Comparación comprensión oral y comprensión lectora

		CL		
		PAR	NoPAR	Total
CO	PAR	2262	7616	9878
		12,9%	43,6%	
	NoPAR	7595	0	7595
		43,5%	0,0%	
Total		9857	7616	17473

CONCLUSIONES

Se analizaron respuestas de una prueba realizada por el Consell Superior d'Avaluació de Catalunya para comprobar las competencias lingüísticas básicas en inglés de los alumnos de sexto curso de educación primaria.

La distribución de puntuaciones, como corresponde a una prueba de competencias básicas, está claramente sesgada hacia las puntuaciones altas.

El 15% de las pautas de respuesta con valores lz^* más bajos han sido consideradas PAR.

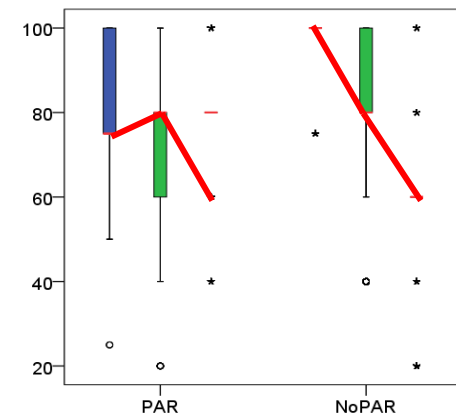
Cerca de un 40% de los suspensos están asociados a PAR. Ese porcentaje va disminuyendo a medida que la nota aumenta. Entre los excelentes el porcentaje se sitúa en el 5% (CO) o en el 8% (CL).

De entre los alumnos con evidencias de PAR en al menos una de las dos pruebas, el 13% presentan patrones atípicos en los dos casos.

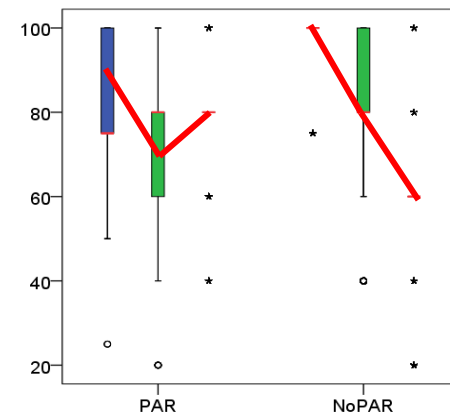
CONCLUSIONES

Los patrones atípicos identificados pueden tener dos consecuencias sobre la interpretación de la puntuación total:

Pueden infravalorar la puntuación total



Pueden sobrevalorar la puntuación total



CONCLUSIONES

Convendría utilizar este tipo de evidencia como fuente de validez de la puntuación total obtenida en este tipo de pruebas.

La acumulación de este tipo de evidencia por parte de un mismo alumno puede ayudar al docente a tomar decisiones.

Acerca de la interpretación de las puntuaciones del alumno (y por ende, de su nivel de competencia).

Acerca de las acciones formativas específicas que puedan ayudar a ese alumno a mejorar su nivel de competencia evaluado.



Creemos que la evaluación de PAR puede constituir una herramienta útil en el ámbito de la evaluación formativa.

REFERENCIAS

Meijer, R. R., and Sijtsma, K. (2001) Methodology review: Evaluating person fit. *Applied Psychological Measurement*, 25(2), 107–135

Snijders, T. B. (2001) Asymptotic null distribution of person fit statistics with estimated person parameter. *Psychometrika*, 66(3), 331–342.

Tendeiro, j. (2015). Package 'PerFit' , versión 1.3.1, 18 de marzo de 2015.
<http://cran.r-project.org/web/packages/PerFit/PerFit.pdf>

Las puntuaciones obtenidas en los tests de conocimientos:
¿son siempre indicadores válidos del aprendizaje?

MUCHAS GRACIAS POR SU ATENCIÓN

eduardo.doval@uab.cat

Este trabajo ha recibido el apoyo económico de la Dirección General de Investigación y Gestión 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.