

## Propuesta para detectar patrones atípicos de respuestas en contextos reales de evaluación

M. Dolors Riba<sup>1</sup>, Eduardo Doval<sup>1</sup>,  
Jordi Renom<sup>2</sup> y Marta Fuentes<sup>3</sup>

<sup>1</sup> *Departament de Psicobiologia i de Metodologia de les Ciències de la Salut  
(UAB),*

<sup>2</sup> *Departament de Psicologia Social i Psicologia Quantitativa  
(UB)*

<sup>3</sup> *Departament de Psicologia Bàsica, Evolutiva i de l'Educació  
(UAB)*

*Dolors.riba@uab.cat*

## INTRODUCCIÓN

---

En las pruebas de evaluación de conocimientos, la existencia de patrones atípicos de respuesta (PAR) pone en duda la validez de las inferencias que puedan realizarse a partir de la puntuación total.

Un PAR es una secuencia de aciertos y errores a los items de una prueba que no se corresponde con el patrón esperado (p.e., respuestas al azar, más respuestas correctas en preguntas difíciles que en preguntas fáciles...)

Para la detección de PAR se han propuesto un gran número de índices (Karabatsos, 2003), algunos paramétricos y otros no paramétricos.

## INTRODUCCIÓN

---

El cálculo de los índices requiere de software específico, y su interpretación no es intuitiva. (Tendeiro, 2016)

El uso de los índices de detección de PAR resulta útil en investigación, pero no es práctico para los docentes que quieren evaluar la presencia de PAR en su grupo clase.

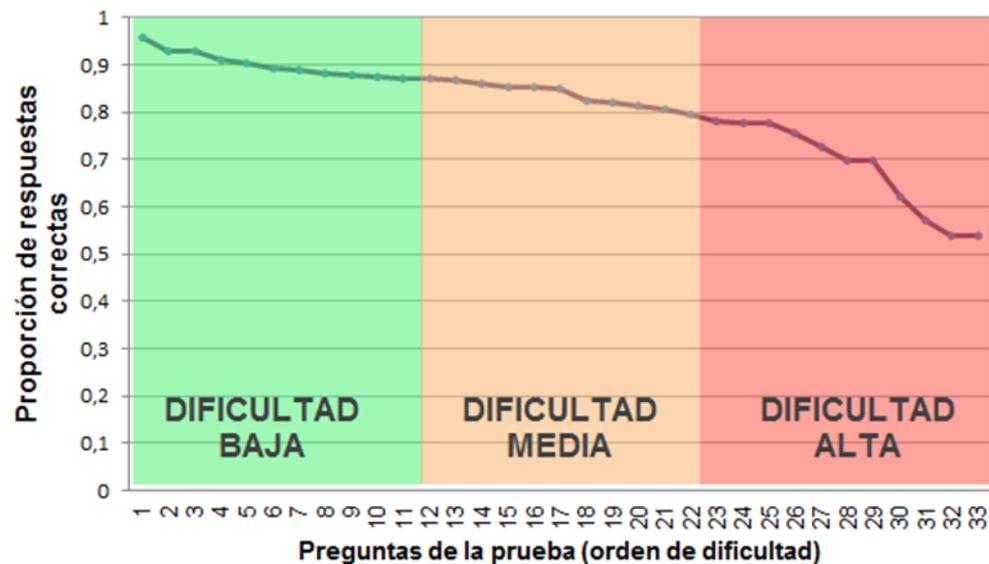
Uno de nuestros intereses ha sido desarrollar una forma de detección de PAR que cumpla esos requisitos de facilidad de cálculo y de interpretación intuitiva.

# INTRODUCCIÓN

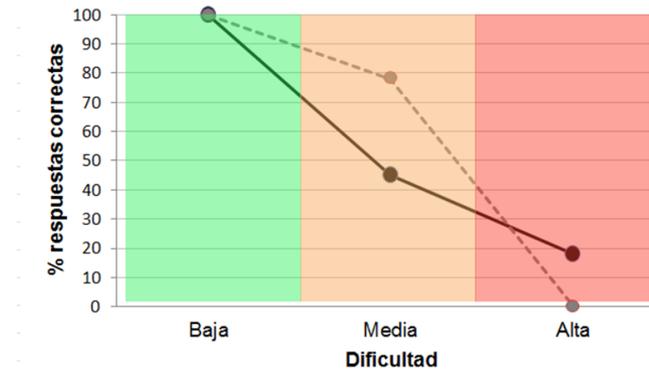
---

El método propuesto consiste en:

- Calcular el grado de dificultad de las preguntas de la prueba a partir de la proporción de respuestas correctas en el grupo de referencia para obtener el perfil de dificultad (de más fácil a más difícil).
- Definir tres bloques de dificultad (baja, media y alta) en función de los centiles (33 y 66).



- Calcular la proporción de respuestas correctas en cada bloque de dificultad (línea continua) y la proporción de respuestas correctas esperada bajo el modelo de Guttman dada la puntuación.



- Calcular la distancia euclídea entre los dos perfiles.

## MÉTODO

---

Se analizan las respuestas de 58.072 cuarto curso de la E.S.O. 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.

## RESULTADOS

### Distancia Euclídea\*MCI tabulación cruzada

			MCI		Total
			No PAR	PAR	
Distancia Euclídea	No PAR	Recuento	37521	8553	46074
		% del total	65,9%	15,0%	80,9%
	PAR	Recuento	8045	2840	10885
		% del total	14,1%	5,0%	19,1%
Total		Recuento	45566	11393	56959
		% del total	80,0%	20,0%	100,0%

Kapa =.074

### Distancia Euclídea\*Iz tabulación cruzada

			Iz		Total
			No PAR	PAR	
Distancia Euclídea	No PAR	Recuento	40764	5310	46074
		% del total	71,6%	9,3%	80,9%
	PAR	Recuento	4809	6076	10885
		% del total	8,4%	10,7%	19,1%
Total		Recuento	45573	11386	56959
		% del total	80,0%	20,0%	100,0%

Kapa =.435

*Kapa (mci-iz) =.586*

# RESULTADOS

Se consideraron válidas las respuestas de 56959 casos (98%).

	Total	Distancia Euc.	MCI	Iz
	41864 (73,5%)	5416 (49,8%)	5361 (47,1%)	4478 (39,3%)
	9156 (16,1%)	1931 (17,7%)	2370 (20,8%)	3009 (26,4%)
	5730 (10,1%)	3361 (30,9%)	3465 (30,4%)	3690 (32,4%)
	209 (0,4%)	177 (1,6%)	197 (1,7%)	209 (1,8%)
	56959 (100%)	10885 (19%)	11393 (20%)	11386 (19,9%)

## CONCLUSIONES

---

El grado de acuerdo entre las clasificaciones de presencia/ausencia de PAR realizada por las distancias euclídeas es bajo con el realizado con índice MCI, y moderado con el índice Iz.

Tomamos el grado de acuerdo entre la distancia euclídea y el índice Iz como una evidencia moderada de la utilidad de dicha distancia para detectar PAR.

Todos los índices detectan un porcentaje elevado de patrones descendentes. Conviene determinar con más profundidad si esos patrones ponen en duda la validez de las puntuaciones

## REFERENCIAS

---

- 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>

**MUCHAS GRACIAS POR SU ATENCIÓN**

*dolors.riba@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 y a Alba Morató su colaboración en este proyecto.



- Identificar 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).

