



Actualidades en Psicología

ISSN: 0258-6444

actualidades.psicologia@ucr.ac.cr

Instituto de Investigaciones Psicológicas

Costa Rica

Delgado González, Ana R.

La medida de la comprensión emocional con el modelo de Rasch

Actualidades en Psicología, vol. 30, núm. 120, 2016, pp. 47-56

Instituto de Investigaciones Psicológicas

San José, Costa Rica

Disponible en: <http://www.redalyc.org/articulo.oa?id=133246313005>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica

Red de Revistas Científicas de América Latina, el Caribe, España y Portugal

Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

## La medida de la comprensión emocional con el modelo de Rasch

### Measuring Emotion Understanding with the Rasch Model

Ana R. Delgado González<sup>1</sup>

Universidad de Salamanca, España

**Resumen.** El objetivo fue construir ítems de comprensión emocional desde la perspectiva del Modelo de Rasch, comparando mediante un experimento aleatorizado los formatos de respuesta verbal/imagen. Participaron 204 sujetos de población general. Se contrastó el efecto del formato (verbal/imagen) y del género en las medidas Rasch de comprensión de emociones de los participantes y se analizó el efecto de la distancia social (cerca/lejos) sobre la dificultad Rasch de los ítems. El formato de respuesta tuvo un efecto significativo sobre las medidas independientemente del género: el formato verbal resultó más fácil que el de imagen. El efecto del género fue significativo, favorable a las mujeres. Los ítems que describían situaciones con receptores cercanos resultaron más fáciles que los que incluían receptores lejanos.

**Palabras clave.** Construcción de ítems, comprensión emocional, diferencias sexuales, formato de respuesta, modelo de Rasch.

**Abstract.** The aim of this study was to construct emotion understanding items from the Rasch Model approach by experimenting through comparison between verbal /image response format. The participants were 204 subjects from a Spanish community sample. A randomized experiment was carried out to test the effect of response format (verbal/image) and participants' gender on the emotion understanding Rasch measurements. The effect on item difficulty of the social distance (close/far) was also contrasted. No interaction effect was found. Response format had a significant effect on measurement regardless of gender: the verbal response format was easier than the image one. There were significant gender differences on emotion understanding favoring women. Items describing situations with close receivers were significantly easier than the items showing far recipients.

**Keywords.** Emotion understanding, item construction, gender differences, Rasch model, response format.

<sup>1</sup>Ana R. Delgado. Facultad de Psicología, Universidad de Salamanca. Dirección Postal: Avenida de la Merced 109-131. 37005 Salamanca, España. E-mail: [adelgado@usal.es](mailto:adelgado@usal.es)

\* Agradecimientos: Al Centro Documental de la Memoria Histórica (CDMH, Salamanca). A Ron Dotsch, Gijs Bijlstra y Oliver Langner que permitieron el acceso a la Radboud Faces Database.



## Introducción

Los tests de aptitudes emocionales tienen cada vez un mayor papel en el ámbito de la ciencia psicológica (Joseph & Newman, 2010; MacCann, Joseph, Newman & Roberts, 2014; Marquez & Delgado, 2012; Mayer, Roberts & Barsade, 2008). Los estudios más recientes de la inteligencia emocional, medida como habilidad, corroboran que las aptitudes de comprensión de emociones indexan un factor diferenciado correspondiente al segundo estrato de la inteligencia general (g), en el que existen diferencias de tamaño medio a favor de las mujeres (MacCann, Joseph, Newman & Roberts, 2014).

El objetivo general de este trabajo fue la construcción de una medida de comprensión emocional desde la perspectiva del modelo de Rasch, un modelo de escalamiento conjunto de personas e ítems cuyas ventajas psicométricas son ya bien conocidas (Engelhard, 2013; Inchausti, Prieto & Delgado, 2014; Prieto & Delgado, 2003; Prieto & Díaz, 2003; Rasch, 1960); contrastando experimentalmente el efecto del formato (verbal/imagen) sobre la dificultad de la tarea. En la tradición de investigación de los métodos mixtos, los enunciados del test se redactaron considerando tanto cómo se teoriza actualmente la emoción (e.g., Tracy & Randles, 2011) como los resultados procedentes de diversos estudios cualitativos que ofrecen información ecológica sobre los receptores y las atribuciones asociadas en países hispanohablantes a emociones poco salientes en el ámbito anglosajón, tales como el desprecio (Delgado, 2009a, 2009b; Delgado & Marquez, 2013a, 2013b). En particular, la distancia psicológica, una variable que se hace evidente en los estudios cualitativos citados, es actualmente objeto de estudio en multitud de investigaciones cognitivo-sociales (Soderberg, Callahan, Kochersberger, Amit & Ledgerwood, 2015).

Contar con ítems que puedan responderse pulsando sobre la imagen de una expresión emocional, en lugar de hacerlo mediante una palabra, podría suponer una gran ventaja tanto desde el punto de vista de la aplicabilidad en determinadas situaciones (e.g., clínicas) como por la posibilidad de comparar resultados con los

de otros países. El acceso a la Radboud Faces Database (Langner, Dotsch, Bijlstra, Wigboldus, Hawk & van Knippenberg, 2010) ha hecho posible la construcción de opciones de respuesta en formato de imagen de alta calidad para la versión *imagen* del test de comprensión de emociones.

En cuanto a las emociones seleccionadas para la construcción de los ítems, una reciente investigación llevada a cabo mediante el modelo de Rasch para contrastar la calidad psicométrica de una tarea de reconocimiento emocional comúnmente empleada, halló que las expresiones faciales de desprecio daban lugar a ítems muy difíciles, seguidos, en este orden, por miedo, asco, rabia y pena; ni los ítems de alegría ni los de sorpresa contribuyeron fructíferamente a la medición en dicho estudio, debido a su extrema facilidad en muestras de población general (Delgado, 2012).

Sin embargo, puesto que incluso los investigadores más escépticos mantienen que el reconocimiento de esas siete expresiones emocionales es universal (e.g., Henrich, Heine and Norenzayan, 2010), se han construido ítems para todas ellas, aunque, por lo anteriormente expuesto, en el diseño del test se ha partido de un mayor número de ítems de desprecio, mientras que la alegría y la sorpresa contaron con un número menor.

## Método

### *Participantes*

Siguiendo las normas éticas, 204 participantes voluntarios de la población general fueron reclutados en un céntrico museo de Salamanca y, tras prestar su consentimiento, completaron los tests en un lugar tranquilo. La muestra se compuso de igual número de varones y mujeres de edades comprendidas entre los 18 y los 65 años (edad promedio = 39.24;  $DE = 12.39$ ) procedentes de las distintas comunidades autónomas de España y con el español como lengua materna.

### *Instrumentos*

El test informatizado se programó en LiveCode 4.6 (2011). Identificación, sexo, edad, consentimiento y opción de respuesta elegida, así como el acierto/

error para cada ítem, se almacenan automáticamente en matrices listas para ser exportadas y analizadas. Los enunciados de los 30 ítems fueron redactados a partir de situaciones sociales prototípicas desarrolladas a partir de las teorías actuales sobre las emociones discretas (también denominadas básicas o universales) y de los resultados de diversos estudios cualitativos llevados a cabo mediante entrevistas y análisis de textos en castellano (Delgado, 2009a, 2009b; Delgado & Marquez, 2013a, 2013b; Tracy & Randles, 2011).

Las instrucciones indican que se trata de situaciones en las que se encuentra un personaje y los sujetos deben seleccionar la palabra (o imagen, dependiendo de la condición experimental) que corresponda a lo que sería más normal que el personaje sintiera en ese caso. Los enunciados de los ítems aparecen en la Tabla 1. Nótese que el personaje al que se atribuye la emoción es femenino en la mitad de los ítems y masculino en el resto; para estos personajes se seleccionaron nombres bisílabos y de muy alta frecuencia en castellano.

Tabla 1

*Comprensión de emociones: características y enunciados de los ítems*

Ítem	Emoción	DP	Enunciado
1	alegría	C	El hermano de Ana le ha regalado un libro que ella estaba deseando leer.
2	sorpresa	C	Inesperadamente, Manuel se ha encontrado con su padre por la calle.
3	pena	C	Ha fallecido el mejor amigo del novio de Ana.
4	alegría	C	La hermana de Manuel le ha regalado un disco que a él le encanta.
5	sorpresa	C	Inesperadamente, Ana se ha encontrado a su madre en una tienda.
6	rabia	C	El mejor amigo de Manuel echó a un mimo del lugar público en el que actuaba.
7	pena	L	Ana acaba de enterarse de que muchos médicos han perdido su empleo.
8	miedo	C	El mejor amigo de Manuel ha recibido amenazas de muerte de una banda de matones.
9	desprecio	C	El novio de Ana se fue de vacaciones cuando ella tenía graves problemas económicos.
10	rabia	C	Manuel piensa que su hermano es demasiado manipulador.
11	desprecio	L	Ana se ha enterado de que unos escritores han escrito un artículo con información falsa contra un actor.
12	pena	C	Manuel se ha enterado de que ha fallecido la mejor amiga de su novia.
13	asco	C	En el hospital, Ana ha visto a cirujanos con manchas de vómito en la bata.
14	miedo	L	Manuel ha leído que unos médicos han sido secuestrados por terroristas que amenazan con matarlos.

Continúa...

Ítem	Emoción	DP	Enunciado
15	<i>rabia</i>	C	La mejor amiga de Ana echó a un músico callejero del lugar público en el que actuaba.
16	<i>desprecio</i>	L	Los médicos no informaron a Manuel de posibles tratamientos cuando estuvo gravemente enfermo.
17	asco	L	En el restaurante, Ana ha visto a cocineros que llevaban pegados trozos de tomate en el delantal.
18	pena	L	Manuel ha leído que muchos periodistas han perdido su empleo.
19	desprecio	L	Ana piensa que algunos políticos locales son unos sinvergüenzas.
20	rabia	C	Manuel piensa que su hermana es demasiado controladora.
21	desprecio	L	Ana se ha enterado de que unos periodistas han escrito un artículo con información falsa contra un profesor.
22	asco	C	En el restaurante, Manuel ha visto a cocineros con manchas de vómito en el delantal.
23	rabia	C	El novio de Ana se opuso a que ella hiciera un curso de formación que deseaba hacer.
24	desprecio	C	La novia de Manuel se fue de viaje con sus amigas cuando él estaba gravemente enfermo.
25	miedo	C	La mejor amiga de Ana ha recibido amenazas de muerte de una banda de matones.
26	asco	L	En el hospital, Manuel ha visto a cirujanos que llevaban pegados trozos de tomate en la bata.
27	desprecio	L	Los abogados del seguro de Ana no le indicaron las opciones que tenía cuando fue estafada.
28	rabia	C	La novia de Manuel se opuso a que él hiciera un viaje que deseaba hacer.
29	miedo	L	Ana ha leído que unos periodistas han sido secuestrados por terroristas que amenazan con matarlos.
30	desprecio	L	Manuel piensa que hay muchos abogados corruptos.

*Nota.* C: cerca; DP: distancia psicológica; L: lejos.

En contextos sociales, las distintas emociones pueden experimentarse por diferentes motivos, es decir, atribuirse a razones diferentes (e.g., reciprocidad: el receptor de la emoción ha hecho algo negativo al emisor; o altruismo: el receptor de la emoción ha hecho algo negativo a un tercero); esto se contempló en las especificaciones para generar variaciones de los prototipos, como puede deducirse al leer los enunciados,

por lo que este aspecto no se incluye en la Tabla 1. Sí se ha señalado en dicha tabla la categorización cerca/lejos relativa a la distancia psicológica con el receptor de la emoción (que, en el caso del asco, diferencia lo corporal de lo externo), ya que esta variable resulta de interés para el análisis de la dificultad de los ítems.

Cada ítem del test cuenta con dos versiones idénticas en cuanto al enunciado, pero cuyas opciones



Figura 1. Ítem 1 del test de comprensión de emociones en ambos formatos.

de respuesta difieren en el formato: verbal o imagen (ver Figura 1). Las etiquetas de los ocho botones de respuesta verbal son: desprecio, rabia, asco, miedo, pena, sorpresa, alegría y neutral.

Los botones de respuesta imagen se construyeron a partir de imágenes seleccionadas de la Radboud Faces Database (Langner et al., 2010, support material); se trata de los planos frontales de las siete expresiones emocionales (más la neutral) de dos modelos; una mujer (la número 16) y un varón (el número 23) elegidos entre los que obtuvieron valores superiores al promedio en expresión del desprecio en el estudio original. El resto de sus expresiones se reconocían con facilidad (por encima del 80%); estos dos modelos eran también los mejores en términos

de expresión genuina e intensidad (Langner et al., 2010, support material).

#### *Procedimiento*

En primer lugar, se preguntaba a los sujetos si deseaban colaborar en una investigación universitaria cuyos datos se analizarían de forma agregada. En caso afirmativo, se preguntaba por su edad, lugar de procedencia y lengua materna para garantizar el cumplimiento de los criterios de inclusión. La aplicación fue individual, mediante un ordenador portátil. La mitad de la muestra recibió la versión verbal del test; la otra mitad completó la versión en formato imagen. La asignación de los participantes a los niveles de la variable manipulada (verbal/imagen) fue aleatoria. La mitad de cada grupo experimental fueron varones y la otra mitad, mujeres.

### Análisis de datos

El modelo de Rasch es un modelo de medición conjunta de personas e ítems en escala *logit*. La probabilidad de que el sujeto  $n$  responda correctamente al ítem  $i$ , se modela con la fórmula  $P_{ni} = \exp(B_n - D_i) / (1 + \exp[B_n - D_i])$ , donde  $B_n$  es el nivel de aptitud del sujeto y  $D_i$  es la dificultad del ítem. Los análisis Rasch se llevaron a cabo con el programa Winsteps 3.80.1 (Linacre, 2013). El ajuste datos-modelo se valora mediante el índice *outfit*, que se calcula a partir de los residuos cuadráticos estandarizados y se distribuye aproximadamente como chi-cuadrado, y el índice *infit*, que se calcula también a partir de los residuos cuadráticos estandarizados, pero ponderados por la función de información. Siguiendo las indicaciones del programa (Linacre, 2013), los valores *infit* / *outfit* superiores a 2.0 degradan la medida; valores entre 1.5 y 2.0 serían improductivos, pero no degradarían las medidas; valores entre .5 y 1.5 serían los adecuados; finalmente, valores inferiores a .5 son menos productivos, pero no degradarían la medida. Para ambos indicadores, se esperan valores promedio para personas e ítems cercanos a 1.0 y con escasa desviación estándar (DE).

La evidencia sobre el ajuste de los datos al modelo incluye la relativa a la unidimensionalidad, que no implica que la ejecución en una tarea se deba a un único proceso psicológico. A partir de las respuestas de los participantes, el modelo de Rasch para datos dicotómicos (Rasch, 1960) permite estimar las medidas de ítems ( $D_i$ ) y personas ( $B_n$ ) en una escala conjunta, en una misma dimensión. Para corroborar el supuesto de unidimensionalidad, el porcentaje de varianza observada que es explicada por esa dimensión debería ser de, al menos, el 20%, sin factores secundarios de importancia (Reckase, 1979); la varianza de los residuos es la varianza no explicada por las medidas Rasch y su análisis sirve para detectar posibles dimensiones secundarias, que han de ser de tamaño suficiente para ser consideradas de interés. En el análisis de componentes de los residuos, cada ítem contribuye una unidad de información (un *eigenvalue*): e.g., un valor de 2 en el primer componente lleva asociada la información (varianza) de 2 ítems (Linacre, 2013).

Además, se llevaron a cabo análisis del funcionamiento diferencial de los ítems (DIF) para contrastar la validez generalizada de los ítems en grupos diferentes. En el caso de dos grupos y respuesta dicotómica, se recomienda el método paramétrico de detección del DIF derivado del modelo de Rasch, que calcula la diferencia estandarizada entre las estimaciones de la dificultad de los ítems en ambos grupos (Smith, 2004). Para cada hipótesis de diferencia, se contrasta la significación estadística mediante el estadístico *t de Welch* utilizando el ajuste Bonferroni (alfa/ número de contrastes). Un contraste significativo de tamaño igual o superior a .50 logit en valor absoluto indica que el DIF es sustancial, además de estadísticamente significativo, por lo que este es el criterio habitual en la interpretación del DIF (Linacre, 2013).

El efecto de la distancia psicológica (cerca/lejos) en la dificultad Rasch de los ítems del test de comprensión de emociones se contrastó con la prueba U de Mann-Whitney. Los efectos del género de los participantes (masculino/femenino) y del formato de respuesta en que se completó el test (verbal/imagen), sobre las puntuaciones Rasch de los sujetos en el test de comprensión de emociones, se contrastaron mediante ANOVA de dos factores. Se trata de un experimento aleatorizado en el que la variable independiente manipulada es el formato de respuesta del ítem, verbal o imagen, y el género de los participantes se ha controlado por inclusión: la mitad de cada grupo experimental son varones y la otra mitad, mujeres.

## Resultados

En el análisis Rasch de las respuestas en formato de respuesta verbal, se encontraron correlaciones negativas ítem-medida para los ítems 16 y 27 (desprecio hacia un receptor lejano). Una vez eliminados ambos ítems, los indicadores de ajuste de ítems y personas fueron adecuados. Para los ítems, el promedio de outfit fue .99 (DE = .10) y el promedio de infit fue 1.00 (DE = .07). Para las personas, el promedio de outfit fue .99 (DE = .26) y el promedio de infit fue 1.00 (DE = .19). Ni en el caso de los ítems ni en el de las personas se hallaron valores de infit/outfit superiores a 2. El ítem



*reliability*, el indicador global de fiabilidad para los ítems fue .95 y el *model person reliability*, el indicador global de fiabilidad para las personas fue .62.

En cuanto al análisis Rasch de las respuestas en formato de respuesta *imagen*, se encontraron correlaciones negativas ítem-medida para los ítems 9 y 24 (desprecio hacia un receptor cercano). Una vez eliminados ambos ítems, los indicadores de ajuste de ítems y personas resultaron adecuados. Para los ítems, el promedio de outfit fue 1.04 ( $DE = .18$ ) y el promedio de infit fue 1.00 ( $DE = .07$ ). En el caso de las personas, el promedio de outfit fue 1.04 ( $DE = .57$ ) y el promedio de infit fue 1.00 ( $DE = .20$ ). Ningún ítem mostró valores de infit/outfit superiores a 2 y solo en el caso de cuatro personas, los valores outfit superaron el valor 2. Item reliability fue .96 y model person reliability fue .57.

Tras haber eliminado los ítems 9, 16, 24 y 27, se analizaron conjuntamente las respuestas procedentes de los 204 sujetos a los restantes 26 ítems del test. Se halló un buen ajuste de las puntuaciones de personas e ítems al modelo de Rasch. Para los ítems, el promedio de outfit fue 1.00 ( $DE = .08$ ) y el de infit fue 1.00 ( $DE = .05$ ). Para las personas, el promedio de outfit fue 1.00 ( $DE = .25$ ) y el de infit fue 1.00 ( $DE = .18$ ). Ningún ítem mostró valores de infit o de outfit superiores a 1.5 y ninguna persona mostró valores de infit o de outfit superiores a 2. El Item reliability fue .97 y model person reliability fue .63.

En la Tabla 2 aparecen los principales resultados del análisis de los ítems. La aptitud promedio de las personas en unidades logit fue -.12, ligeramente por debajo del promedio de dificultad de los ítems (donde se sitúa convencionalmente el cero de la escala). El rango de aptitud de la muestra se encuentra entre -2.36 y 1.96.

El porcentaje de varianza explicado por el modelo fue 22.5%, un valor suficiente según el criterio habitual (Reckase, 1979), aunque el análisis de componentes principales de los residuos mostró que la varianza no explicada en el primer contraste era de 2.6, ligeramente por encima de lo recomendado, que es 2. Valorando estos datos en conjunto con los indicadores infit y outfit, el ajuste al modelo de Rasch puede considerarse aceptable.

Tabla 2

*Test de comprensión de emociones (eliminados los ítems 9, 16, 24 y 27): resultados del análisis Rasch*

Ítem	Rasch $d_i$	E r r o r Estándar	Outfit	Infit
1	-1.49	.18	1.04	1.03
2	.22	.15	.98	.98
3	-.52	.15	.83	.86
4	-1.52	.18	.99	.99
5	-.14	.15	1.00	1.00
6	.69	.16	1.09	1.03
7	1.45	.18	.97	1.01
8	-.09	.15	.94	.95
10	1.11	.17	1.10	1.05
11	.92	.16	1.07	1.03
12	-1.40	.17	.96	.96
13	-.99	.16	1.07	1.07
14	.64	.16	1.08	1.03
15	.54	.15	1.08	1.07
17	-.59	.15	.99	.98
18	.69	.16	.96	.94
19	1.00	.17	.87	.92
20	.86	.16	.98	1.01
21	1.08	.17	1.06	1.05
22	-1.90	.19	.86	.97
23	-.20	.15	1.05	1.05
25	-.72	.15	.90	.93
26	-.56	.15	1.04	1.02
28	.17	.15	1.00	1.02
29	.31	.15	.96	.96
30	.47	.15	1.16	1.07



Por otra parte, ningún ítem mostró DIF asociado al género. Los análisis del DIF asociado al grupo de respuesta (verbal/imagen) mostraron que tres ítems favorecían al grupo que respondió en formato verbal (i3, i7, i18) y otros tres favorecían al grupo de respuesta en formato imagen (i13, i15, i23). Puesto que el número de ítems que favorece a cada grupo es el mismo y no es elevado proseguir con el análisis conjunto de las versiones de los 26 ítems de comprensión de emociones es justificable, por cancelación del DIF (Roznowski y Reith, 1999).

Los resultados del ANOVA factorial,  $F_{(3, 200)} = 15.25$ ,  $p = .001$ , indican que tanto el formato de respuesta (verbal/imagen),  $F_{(1, 200)} = 26.34$ ,  $p = .001$ ,  $d = -.69$ , como el género de los participantes,  $F_{(1, 200)} = 17.72$ ,  $p = .00$ ,  $d = -.56$ , tuvieron efectos significativos y por encima del considerado como tamaño medio sobre las medidas Rasch de comprensión de emociones de los sujetos. El efecto de la interacción no resultó estadísticamente significativo  $F_{(1, 200)} = 1.69$ ,  $p = .20$ . El formato de respuesta verbal facilitó la tarea, como indica la mayor puntuación promedio del grupo que respondió en dicho formato,  $M(\text{verbal}) = .12$ ,  $DE = .72$ ;  $M(\text{imagen}) = -.36$ ,  $DE = .68$ . En cuanto a la variable género, las mujeres obtuvieron una media superior a la de los varones,  $M(\text{femenino}) = .08$ ,  $DE = .74$ ;  $M(\text{masculino}) = -.32$ ,  $DE = .69$ .

Es de resaltar que, en este caso, los resultados son muy similares a los que se obtienen analizando mediante ANOVA las puntuaciones directas obtenidas en el experimento, que es el procedimiento habitual de análisis; sin embargo, dicho procedimiento no garantiza que la variable dependiente esté medida en una escala de intervalo, con una unidad constante, algo que aquí se ha contrastado mediante el modelo de Rasch. Una vez realizado este contraste, puede defenderse la interpretación cuantitativa, no meramente ordinal o cualitativa, de las puntuaciones. El mapa del constructo, en el que se representan conjuntamente la aptitud de los participantes (los más aptos arriba) y la dificultad de los ítems (los más difíciles arriba) puede verse en la Tabla 3.

Por último, se contrastó el efecto de la distancia psicológica (cerca/lejos) del receptor de la emoción en

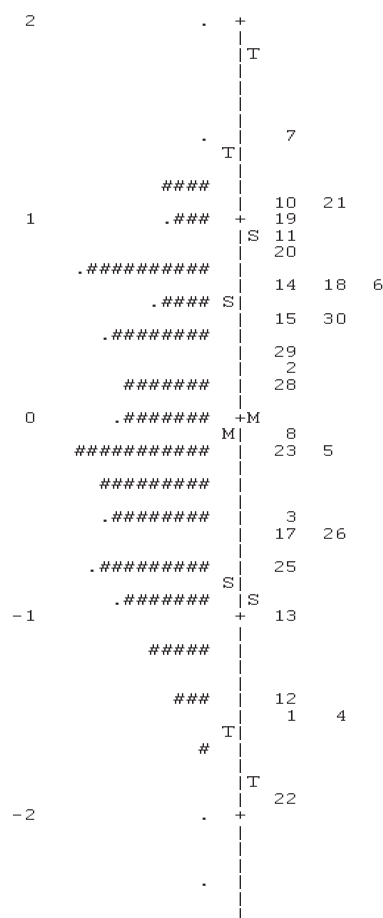
la dificultad Rasch de los ítems del test de comprensión de emociones,  $U \text{ de Mann-Whitney} = 36.50$ ,  $z = -2.29$ ,  $p = .022$ . Los ítems que describen situaciones con receptores cercanos resultaron más fáciles,  $M(\text{cerca}) = -.34$ ,  $DE = .93$ ;  $M(\text{lejos}) = .54$ ,  $DE = .67$ ,  $d = -1.09$ .

## Discusión

Se han puesto a prueba, con ayuda del modelo de Rasch, las propiedades psicométricas de las

Tabla 3

Test de comprensión de emociones (eliminados los ítems 9, 16, 24 y 27):  
representación conjunta de personas e ítems



Nota. La almohadilla representa a 2 personas y el punto a 1. DT: Desviación Típica; M: media; S: 1 DT; T: 2DT.

puntuaciones de un test de aptitud emocional que ha sido desarrollado tomando en consideración aspectos teóricos y ecológicos de emociones con distinto peso en nuestro entorno y en el ámbito anglosajón, tales como la rabia y el desprecio (Delgado, 2009a, 2009b; Delgado & Marquez, 2013a, 2013b). Además de construir los enunciados de los ítems contando con información procedente de estudios cualitativos de entrevistas y textos en español, se ha puesto a prueba el efecto del formato de respuesta por medio de un verdadero experimento.

En la validación inicial, se ha contado con un gran número de ítems relativos al desprecio por tratarse de una emoción particularmente saliente en España (Delgado, 2009a). Por otra parte, se había hallado que los ítems de desprecio resultaban más difíciles que los del resto de las denominadas emociones básicas, por lo que sirven para evitar el efecto techo (Delgado, 2012).

De los treinta ítems iniciales, ocho correspondían al desprecio y cuatro de ellos fueron eliminados, dos a partir del análisis de la versión verbal y otros dos a partir del análisis de la versión de imagen. En relación con el escenario descrito en el enunciado, puede verse en la Tabla 1 que los cuatro ítems corresponden a situaciones en las que el desprecio se puede describir como “recíproco”, pero en dos de los ítems se trata de una situación en la que la distancia psicológica con el receptor es grande; mientras que, en los dos restantes es pequeña. Se trata de un receptor cercano. Puesto que aún quedan cuatro ítems de desprecio en la versión de 26 ítems, esta emoción está suficientemente representada en el test.

En cuanto a las diferencias asociadas al sexo en las aptitudes de procesamiento de emociones, es habitual hallar que las mujeres obtienen mejores puntuaciones (Joseph & Newman, 2010; MacCann, Joseph, Newman & Roberts, 2014), y el hecho de que se hayan replicado dichas diferencias en este estudio es un indicador de validez. Además, la ausencia de efecto de la interacción (género x formato de respuesta) permite una clara interpretación del efecto del formato de respuesta (verbal/imagen) en las medidas Rasch de los sujetos: los ítems con formato

de respuesta verbal resultan más fáciles que los ítems con formato imagen.

Si seguimos la sugerencia de Newton y Shaw (2013) de favorecer el criterio de calidad en relación a distintos objetivos (políticos, de toma de decisiones o de medida), se puede concluir que la mayor parte de los ítems de comprensión de emociones generados podrían ser ya empleados como componentes de un proceso de medida o decisión de mayor nivel. Añadir nuevos ítems, similares a los mejores de entre los ya empleados en esta investigación, permitirá mejorar la fiabilidad en futuras versiones. Llevar a cabo estudios de validación predictiva en los ámbitos de interés es el siguiente paso para contrastar la calidad de las medidas.

Considerando la relevancia de la inteligencia emocional en las modernas economías de servicios, así como la evidencia que corrobora que la inteligencia emocional es uno de los factores del segundo estrato de g, ya se recomienda la inclusión de ítems similares a los aquí desarrollados en los futuros tests de inteligencia (MacCann, Joseph, Newman & Roberts, 2014).

Dada la escasez de medidas de este tipo, parece deseable trabajar en proceso de construcción y validación de pruebas de aptitudes emocionales por medio de modelos psicométricos avanzados, como el modelo de Rasch, pruebas que estén informadas por estudios del idioma en el que van a emplearse, como las aquí descritas, ya que existen variantes emocionales poco empleadas en inglés que pueden ser, sin embargo, clínicamente importantes.

## Referencias

- Delgado, A. R. (2012). Using the Rasch Model to test the psychometric quality of emotion recognition items. *Journal of the International Neuropsychological Society*, 18 (S1), 69.
- Delgado, A. R. (2009a). Spanish basic emotion words are consistently ordered. *Quality & Quantity*, 43, 509-517.
- Delgado, A. R., (2009b). Social robots, moral emotions. Proceedings of the 11th International Conference on Enterprise Information Systems, vol. AIDSS (pp. 263-270). ICEIS.

- Delgado, A. R., & Máarquez, M. G. (2013a). Intracategorical varieties of contempt in a heterogeneous sample. *Universitas Psychologica*, 12, 357-362.
- Delgado, A. R., & Máarquez, M. G. (2013b). *Social robots, cross-cultural differences*. Proceedings of the 15th International Conference on Enterprise Information Systems, vol. 3 (pp. 109-112). INSTICC.
- Engelhard, G. (2013). *Invariant measurement: Using Rasch Models in the social, behavioral and health sciences*. New York, NY: Routledge.
- Henrich, J. Heine, S.J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33, 61-135.
- Inchausti, F., Prieto, G., & Delgado, A. R. (2014). Análisis Rasch de la versión española de la escala Mindful Attention Awareness Scale en una muestra clínica. *Revista de Psiquiatría y Salud Mental*, 7, 32-41.
- Joseph, D. L., & Newman, D. A. (2010). Emotional intelligence: an integrative meta-analysis and cascading model. *Journal of Applied Psychology*, 95, 54-78.
- Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud faces database. *Cognition and Emotion*, 24, 1377-1388. DOI:10.1080/02699930903485076
- Linacre, J. M. (2013). *Winsteps Rasch measurement computer program, version 3.80.1*. Chicago: Winsteps.com.
- LiveCode 4.6 [computer software]. (2011). Retrieved from <http://runrev.com/>
- MacCann, C., Joseph, D. L., Newman, D. A., & Roberts, R. D. (2014). Emotional intelligence is a second stratum factor of intelligence: evidence from hierarchical and bifactor models. *Emotion*, 14, 358-74.
- Marquez, M. G., & Delgado, A. R. (2012). Revisión de las medidas de reconocimiento y expresión de emociones. *Anales de Psicología*, 28, 978-985.
- Mayer, J. D., Roberts, R. D., & Barsade, S. G. (2008). Human abilities: emotional intelligence. *Annual Review of Psychology*, 59, 507-536.
- Newton, P. E., & Shaw, S. D. (2013). Standards for talking and thinking about validity. *Psychological Methods*, 18, 301-19.
- Prieto, G., & Delgado, A. R. (2003). Análisis de un test mediante el modelo de Rasch. [Rasch-modelling a test]. *Psicothema*, 15, 94-100.
- Prieto, G., & Díaz Velasco, A. (2003). Uso del modelo de Rasch para poner en la misma escala las puntuaciones de distintos tests. *Actualidades en Psicología*, 19, 5-23.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen, Denmark: Danish Institute for Educational Research.
- Reckase, M. D. (1979). Unifactor latent trait models applied to multifactor tests: results and implications. *Journal of Educational Statistics*, 4, 207-230.
- Roznowski, M., & y Reith, J. (1999). Examining the measurement quality of tests containing differentially functioning items: do biased items result in poor measurement? *Educational and Psychological Measurement*, 59, 248-69.
- Smith, R. M. (2004). Detecting item bias with the Rasch model. *Journal of Applied Measurement*, 5, 430-449.
- Soderberg, C. K., Callahan, S. P., Kochersberger, A., Amit, E., & Ledgerwood, A. (2015). The effect of psychological distance on abstraction: A meta-analysis of construal level theory. *Psychological Bulletin*, 141, 525-548.
- Tracy, J. L., & Randles, D. (2011). Four models of basic emotions: a review of Ekman and Cordaro, Izard, Levenson, and Panksepp and Watt. *Emotion Review*, 3, 397-405.

Recibido: 19 de octubre de 2015

Aceptado: 10 de marzo de 2016