

## Revisión de las medidas de reconocimiento y expresión de emociones

Margarita G. Marquez\* y Ana R. Delgado

Universidad de Salamanca (España)

**Resumen:** El objetivo de esta revisión sistemática fue analizar las características psicométricas de los instrumentos de medida del reconocimiento y la expresión de emociones. El trabajo se realizó en dos fases. Las categorías empleadas en la primera fase, deductiva, fueron las palabras clave que llevaron a la selección de 721 resúmenes de artículos procedentes de las bases de datos de Scopus y PsycInfo. El proceso selectivo de la segunda fase fue inductivo, utilizando categorías construidas a partir del análisis de los resúmenes de la primera fase. Los instrumentos contenidos en los 36 artículos de investigación seleccionados en esta segunda fase se codificaron en función de la muestra, el tipo de evaluación y el análisis psicométrico efectuado en cada caso. Los resultados muestran que la mayor parte de las medidas derivan de un solo test y que no se encuentran instrumentos contruidos a partir de métodos modernos como el Modelo de Rasch. Los avances de la Neuropsicología afectiva no se han acompañado del desarrollo de instrumentos estandarizados de evaluación y medida que permitan cuantificar adecuadamente las diferencias individuales.

**Palabras clave:** medición; modelo de Rasch; neuropsicología afectiva; expresión de emociones; reconocimiento de emociones; revisión.

**Title:** A review of emotion recognition and expression measures.

**Abstract:** A systematic review was carried out with the objective of learning the Psychometric properties of measures of emotion recognition and expression. They were selected in two steps: first, a deductive one in which 721 abstracts from Scopus and PsycInfo were analyzed; and second, the inductive one, by using the *ad hoc* categories from the previous analysis. The instruments from the 36 selected papers were coded by sample, assessment type, and psychometric analyses. Results show that most of the measures come from just one test and that modern test models such as Rasch-based ones are never found. It is concluded that Affective Neuropsychology advances have not been accompanied by the construction of standardized measurement instruments to properly quantify individual differences.

**Key words:** measurement; Rasch model; affective neuropsychology; emotion expression; emotion recognition; review.

### Introducción

El procesamiento de las expresiones emocionales es fundamental para la adaptación psicológica y conductual de los seres humanos; está implicado en aspectos de la esfera funcional del individuo tan relevantes como el aprendizaje o la toma de decisiones (Nakhutina, Borod y Zgaljardic, 2006). Uno de los aspectos clave de la interacción entre las funciones que se consideran cognitivas y las emocionales –los modelos tradicionales las consideran como claramente diferenciadas, aunque la evidencia empírica está diluyendo los límites entre ambas– es que puede proporcionar información sobre posibles marcadores fenotípicos relacionados con ciertos desórdenes psiquiátricos y neuropsicológicos como la esquizofrenia o la demencia tipo Alzheimer (Williams et al., 2009). De ahí que, en los últimos años, se haya empezado a prestar atención a dicha interacción tanto en condiciones normales como en situaciones de déficit (Mathersul et al., 2009).

Los procesos neuropsicológicos que subyacen a lo afectivo dependen de dos redes neurales muy relacionadas (Bozikas et al., 2009). La primera es la red ventral compuesta por la amígdala, la ínsula, el estriado ventral, regiones ventrales del cíngulo anterior y el córtex prefrontal ventrolateral y orbitofrontal. La segunda red es el sistema dorsal y está formada por el hipocampo y por regiones dorsales del giro cíngulo anterior y por diversas áreas del córtex prefrontal. Mientras que la red anterior es importante para la identificación de estímulos ambientales con significado emocional, la red dorsal permite el desarrollo de funciones ejecutivas y, a

su vez, la regulación de estados afectivos. Según Philips, Drevets, Raugh y Lane (2003) los trastornos psicopatológicos que cursan con alteraciones en las redes neurales a nivel estructural y/o funcional podrían estar asociados a déficits en el procesamiento afectivo.

Aunque el inicio del interés científico por la comunicación no verbal de tipo afectivo se originó con la investigación de Darwin (1872) sobre el impacto de las expresiones faciales y corporales en el lenguaje verbal, no fue hasta los años 70 del siglo XX cuando empezó a generarse un cuerpo articulado de investigación sobre la expresión gestual de emociones. En la primera década del siglo actual, numerosas investigaciones neuropsicológicas han utilizado expresiones faciales como estímulos a procesar en estudios con técnicas de neuroimagen. A su vez, los resultados obtenidos en este tipo de tareas se han empleado para establecer correlaciones con medidas neuroanatómicas o con el efecto ocasionado por daños a nivel cerebral de diversa consideración (Gur et al., 2002).

Si bien es notable el avance de la investigación neuropsicológica acerca de la comunicación no verbal y el procesamiento emocional, es también evidente que se ha obviado hasta cierto punto el tema de las diferencias individuales en estas variables (reconocimiento o expresión de expresiones faciales) así como su medida estandarizada (Suzuki, Hoshino y Shigamasu, 2006); no ha habido excesivo interés por desarrollar instrumentos de medida con validez de constructo capaces de diagnosticar las diferencias individuales (Bänzinger, Grandjean, y Scherer, 2009). La evaluación neuropsicológica del reconocimiento o de la expresión emocional se realiza actualmente con expresiones y escenarios prototípicos extraídos en su mayoría de pruebas estandarizadas ya existentes como el *Pictures of Facial Affect* de Ekman y Friesen (1976) o las *Japanese and Caucasian Facial Expressions of Emotion*

\* Dirección para correspondencia [Correspondence address]: Margarita Gómez Márquez. Dpto. Psicología básica, Psicobiología y Metodología de las Ciencias del Comportamiento. Avda. de la Merced 109-131, 37005 Salamanca (España). E-mail: [mgmarquez@usal.es](mailto:mgmarquez@usal.es)

de Matsumoto y Ekman (1988). Según Suzuki et al., (2006), la mayor parte de las pruebas estandarizadas de reconocimiento o expresión emocional mezclan distintas expresiones faciales con niveles muy variables de dificultad. La corriente psicométrica de la Teoría clásica de los tests (TCT) – la más utilizada actualmente en el análisis de los resultados – calcula la suma de las puntuaciones de los individuos sin tener en cuenta esas variaciones en dificultad.

Los modelos de medición conjunta como el Modelo de Rasch (1960), englobados estadísticamente en la Teoría de respuesta al ítem o TRI, pueden solucionar el inconveniente metodológico arriba mencionado y otros problemas asociados a la TCT. En el modelo propuesto por Rasch, el más parsimonioso, el nivel de la persona en el atributo y la dificultad del ítem son los únicos factores que determinan la probabilidad de que una respuesta sea correcta. Si los datos ajustan suficientemente, las puntuaciones tienen propiedades de intervalo y esto garantiza que las diferencias halladas entre los sujetos sean equivalentes a lo largo de todo el continuo. Además, los parámetros estimados para los ítems de la escala son independientes de la muestra de sujetos que resuelven el test y, a su vez, los parámetros estimados para dichos sujetos son independientes de la muestra de ítems seleccionada. Este modelo permite realizar estimaciones de la fiabilidad de los distintos parámetros no sólo para el grupo (como ocurre con la TCT) sino también para cada individuo y facilita el análisis del Funcionamiento Diferencial de los ítems asociado a variables que el investigador pueda considerar relevantes (como, por ejemplo, el origen cultural del observador en el reconocimiento de ciertas emociones). Por tanto, sería relevante revisar los tests de reconocimiento o expresión emocional partiendo de modelos psicométricos avanzados como el modelo de Rasch. Este modelo ha sido suficientemente divulgado en otros contextos como el de la medición del rendimiento académico (Prieto y Delgado, 2007) o en distintas áreas de la psicología clínica como son la evaluación de los trastornos del estado de ánimo y las adicciones (Cadavid, Delgado y Prieto, 2007; Lopes, Delgado, Prieto, Gamito y Trigo, 2010).

El objetivo principal de esta revisión sistemática fue realizar un análisis de los instrumentos de evaluación del reconocimiento y expresión emocional creados hasta la fecha con especial atención a los aspectos psicométricos y, en particular, a la evidencia neuropsicológica empleada en su validación.

## Método

### Corpus

La revisión se llevó a cabo en dos fases. La primera fase de la revisión dio lugar a 854 resúmenes de artículos procedentes de la base de datos de Scopus y 49 resúmenes procedentes de PsycInfo, que, una vez eliminadas las repeticiones, se redujeron a un total de 721. La segunda fase se llevó a cabo sobre 36 artículos de investigación seleccionados de la lis-

ta de resúmenes de la primera fase. Los artículos contenían 33 instrumentos de medida de reconocimiento y expresión de emociones (32 de reconocimiento emocional y solo uno de expresión emocional) presentados por vía sensorial visual y/o acústica. De ellos, 18 eran investigaciones empíricas de tipo correlacional y 18 correspondían a análisis psicométricos de los instrumentos de medida (de los cuales 2 empleaban la técnica del análisis factorial). En el Anexo 1 puede verse un listado de estos tests y de las referencias correspondientes.

### Procedimiento

En primer lugar se seleccionaron los resúmenes de los artículos publicados sobre el reconocimiento y/o la expresión emocional sin incluir ningún límite temporal. Se recurrió a Scopus por tratarse de la base de datos de mayor cobertura en Ciencias de la salud, Medicina, Tecnología y Ciencias Sociales. La búsqueda se efectuó los días 2 y 9 de marzo del 2010 y se completó con las referencias halladas en PsycInfo en esas mismas fechas. En esta primera fase se trabajó deductivamente, utilizando palabras clave estándar: *emotion processing* (EP), *emotion discrimination* (ED), *emotion recognition* (ER), *facial expression recognition* (FER) y *emotional intelligence* combinadas con *testing*, *validity*, *Rasch Model*, *Item Response Theory* (IRT), *reliability* e *individual differences* (ID). Se solicitó que cualquiera de las combinaciones de la búsqueda estuviera presente en el título del artículo, el resumen o las palabras clave del artículo. Ciertas combinaciones no dieron resultados en ninguna de las bases de datos. Éstas fueron las que incluían los términos *Rasch Model* o *IRT* (a excepción de la combinación de *Emotion recognition* con *IRT* que dio como resultado un solo artículo).

Los 36 artículos de la segunda fase de la revisión se seleccionaron tras (1) eliminar los solapamientos existentes entre las dos bases de datos y entre las distintas búsquedas y (2) codificar los 721 resúmenes resultantes mediante categorías temáticas de tal modo que se incluyeran en la segunda fase sólo aquellos artículos que, dentro de cada combinación de búsqueda, abordaran la cuestión específica de la medida estandarizada del reconocimiento y/o la expresión emocional. Se excluyeron los que trataban sobre temas asociados como el procesamiento emocional general o el facial (sin expresión emocional). Los instrumentos contenidos en los 36 artículos de investigación así seleccionados se codificaron mediante sistemas de categorías *ad hoc* relativos a la muestra, el tipo de evaluación y el análisis psicométrico efectuado.

Se estudiaron varios atributos de las muestras: la edad (infancia, adolescencia, edad adulta o vejez), la naturaleza de las mismas (normal o patológica), el tipo de déficit, en su caso, y la existencia o no de un grupo de control a la hora de establecer comparaciones en la variable de reconocimiento o expresión de emociones.

Con respecto al tipo de escalas se consideró el número de tests de reconocimiento emocional recogidos en cada uno de los artículos así como la novedad o la existencia previa de

dichas medidas. Como el análisis de la información estaba más centrado en la naturaleza de las medidas que en los estudios donde éstas se utilizaban, se codificaron características importantes de cada instrumento relacionadas con su naturaleza (medidas de reconocimiento o expresión emocional), las variaciones estímulares realizadas en cada una de las pruebas (cambios en el tipo de emoción, variaciones de intensidad) y la respuesta requerida por parte del sujeto (detección de la emoción y/o detección de la intensidad). Asimismo, se contabilizó el número de tests que recogía cada una de las emociones básicas de Ekman (alegría, asco, desprecio, miedo, pena, rabia, sorpresa) y la vía sensorial por la que se emitían (vía visual o reconocimiento de emociones faciales frente a reconocimiento de aspectos prosódicos del lenguaje).

En cuanto a los análisis psicométricos, se examinó si los análisis de las puntuaciones se habían llevado a cabo con el Modelo de Rasch o con cualquiera de las otras opciones englobadas estadísticamente dentro de la TRI y se analizaron los procedimientos de evaluación de la fiabilidad y la validez, así como la sensibilidad de los tests.

**Tabla 1.** Frecuencia de estudios por fuente y búsqueda.

|   | Scopus     | Psycinfo  | Reconocimiento o expresión emocional | Medidas estándar |
|---|------------|-----------|--------------------------------------|------------------|
| Emotion processing y Testing                            | 104        | 4         | 18                                   | 10               |
| Emotion processing y Validity                           | 96         | 2         | 13                                   | 1                |
| Emotion processing y Reliability                        | 38         | 0         | 8                                    | 2                |
| Emotion recognition e Item response theory              | 0          | 1         | 1                                    | 1                |
| Facial expression recognition y Testing                 | 46         | 3         | 23                                   | 5                |
| Facial expression recognition y Validity                | 23         | 0         | 13                                   | 6                |
| Facial expression recognition y Reliability             | 15         | 1         | 7                                    | 1                |
| Facial expression recognition e Item response theory    | 1          | 0         | 0                                    | 0                |
| Facial expression recognition e individual differences. | 141        | 1         | 62                                   | 10               |
| Emotion processing e individual differences             | 236        | 9         | 21                                   | 0                |
| Emotional intelligence e Item response theory           | 0          | 0         | 0                                    | 0                |
| <b>Total</b>  | <b>700</b> | <b>21</b> | <b>166</b>                           | <b>36</b>        |

## Resultados

En la Tabla 1 se muestran los resultados de la primera fase de la revisión así como del proceso selectivo al que fue sometido el corpus de resúmenes mediante categorías temáticas. Se observa que el número de artículos que emplean medidas estandarizadas de reconocimiento y expresión de emociones -36 artículos- es bastante reducido en comparación al volumen de estudios que realizan dichas mediciones por medio de estímulos que no han sido evaluados psicométricamente. En concreto, 41 de estos últimos realizan mediciones relevantes neuropsicológicamente, tales como la medición del procesamiento emocional en pacientes con daño cerebral. Es importante destacar dentro de la Tabla 1 que, una vez eliminados los solapamientos, solo se extrajo de Psycinfo un resumen en la búsqueda que combinaba *emotion processing y testing* y tres resúmenes para la combinación de *facial expression recognition y testing*.

Los resultados de la segunda fase toman en consideración varios aspectos de los artículos y de los instrumentos de medida, que se codificaron mediante sistemas de categorías *ad hoc* procedentes del análisis cualitativo inicial de los 36 artículos.

El primer sistema de categorías se refiere a la *naturaleza de las muestras*, en concreto a su edad y naturaleza. Los resultados del análisis muestran que, en conjunto, las muestras son bastante heterogéneas si bien predominan los estudios que utilizan muestras patológicas de edad adulta, como puede verse en la Tabla 2.

**Tabla 2.** Frecuencia de estudios por grupo de edad y naturaleza de la muestra.

| Edad                  | Normales  | Patológicas | Total     |
|-----------------------|-----------|-------------|-----------|
| Infancia              | 1         | 2           | 3         |
| Adolescencia          | 2         | 1           | 3         |
| Edad adulta.          | 9         | 16          | 25        |
| Vejez                 | 0         | 2           | 2         |
| Varios grupos de edad | 2         | 1           | 3         |
| <b>Total</b>          | <b>14</b> | <b>22</b>   | <b>36</b> |

De los 22 artículos con muestras patológicas, 15 tienen objetivos correlacionales y 7 realizan revisiones psicométricas de distintas medidas estandarizadas. En relación al déficit específico de las muestras patológicas de este grupo de estudios, predominan aquellos que analizan el reconocimiento emocional de muestras con trastornos como la esquizofrenia (6 artículos), el daño cerebral (4 artículos), la demencia tipo Alzheimer (2 artículos), los trastornos afectivos (3 artículos), el trastorno autista (2 artículos) y los accidentes cerebrovasculares (2 artículos). Hay artículos que, de forma individual, aplican medidas de reconocimiento emocional a muestras con trastorno bipolar, trastorno obsesivo-compulsivo o déficit sensoriales como la baja visión. En 13 de los 22 estudios llevados a cabo con muestras patológicas se cuenta con una muestra control a la hora de medir el reconocimiento o la expresión emocional.

En referencia a los 14 estudios con muestras normales, destaca la cantidad de artículos dedicados al análisis psicométrico de medidas de reconocimiento emocional (9 artículos) en comparación a los análisis factoriales (2 artículos) y a aquéllos de tipo correlacional (3 artículos). Tanto para este tipo de estudios como para los realizados con sujetos patológicos ya mencionados anteriormente, destaca la cantidad de sujetos de edad adulta en comparación a los niños, adolescentes y ancianos.

En cuanto al *tipo de estímulos* empleados en cada test, siete de los 36 estudios aquí revisados utilizan estímulos procedentes de más de un test para establecer medidas de reconocimiento emocional. En total hay 33 tests distintos (32 de reconocimiento y uno de expresión emocional). Trece de los 32 tests de reconocimiento aplican derivaciones de los instrumentos desarrollados por Ekman y Friesen (1976) o Matsumoto y Ekman (1989). Estos dos instrumentos son el *Affect Processing Test* y el *JACBART*. Con respecto a los tests con estímulos distintos a los propuestos por Ekman, la mayor parte de ellos son medidas en fase experimental que se ponen a prueba en un único estudio. Solo tres de ellos se aplican en más de un artículo: el test *DANVA* de Nowicki y Duke (1994) que aparece desarrollado en su versión inicial y en una versión modificada (*DANVA-2*) de Nowicki y Carton (1993), el *Facial discrimination task* de Erwin et al. (1992) y el *Florida Affect Battery* de Bower, Blender y Heilman (1989, citado en Tonks, Williams, Frampton, Yates y Slater (2007).

Con respecto al *tipo de evaluación* realizada en los tests, se puede apreciar en la Tabla 3 que existen instrumentos que evalúan la capacidad del sujeto para identificar distintas emociones presentadas, para indicar la intensidad con la que esas emociones se presentan—por ejemplo, marcándolo en una escala tipo Likert o bien seleccionando, de una pareja de caras con las mismas emociones, aquella que la expresa con mayor intensidad— y otros que integran ambos tipos de respuesta. En todos los instrumentos ha habido previamente una manipulación del tipo de emoción presentada o bien del tipo y la intensidad con que se expresan. Se puede observar en la Tabla 3 que la mayoría se centra en la manipulación y medición del reconocimiento de tipos de emoción.

**Tabla 3.** Frecuencia de tests de reconocimiento emocional para cada tipo de variación en la emoción y para cada tipo de respuesta requerida.

| Tipo de respuesta.   | Variaciones en el tipo de emoción. | Variaciones en tipo e intensidad. | Total |
|--|------------------------------------|-----------------------------------|-------|
| Identificación de la emoción.                              | 23                                 | 4                                 | 27    |
| Selección de la intensidad de una emoción.                 | 1                                  | 2                                 | 3     |
| Identificación de la emoción y selección de su intensidad. | 1                                  | 1                                 | 2     |
| Total  | 25                                 | 7                                 | 32    |

La tabla de contingencia de estas dos variables presenta 5 casillas con frecuencia esperada inferior a 5 – un 83.3 % de las casillas— por lo que no es recomendable utilizar un con-

traste estadístico para poner a prueba la asociación entre variables. En todo caso, se puede apreciar que la mayoría de las medidas de reconocimiento emocional hace variar el tipo de emoción, aunque en ocasiones se haga variar la intensidad de la emoción además del tipo. La respuesta que se exige por parte del sujeto es la identificación del tipo de emoción expresada en la mayoría de los casos; solo un número reducido de instrumentos solicita una evaluación de la intensidad con que se manifiesta la emoción en cuestión.

En referencia a las variables *tipo de emoción* y *vía sensorial de presentación* en las distintas medidas, la más expuesta de todas las emociones básicas de Ekman es la alegría y la menos utilizada es el desprecio (ver Tabla 4).

**Tabla 4.** Frecuencia de tests por emoción y vía de presentación.

|           | Vía visual | Vía acústica (prosodia) | Ambas vías | Total |
|-----------|------------|-------------------------|------------|-------|
| Alegría   | 24         | 2                       | 6          | 32    |
| Asco      | 19         | 0                       | 1          | 20    |
| Desprecio | 3          | 0                       | 0          | 3     |
| Miedo     | 17         | 1                       | 5          | 23    |
| Pena      | 23         | 2                       | 6          | 31    |
| Rabia     | 20         | 2                       | 6          | 28    |
| Sorpresa  | 18         | 1                       | 2          | 21    |
| Neutra    | 5          | 2                       | 3          | 10    |

Con respecto al tipo de *vía sensorial* de exposición de los estímulos, se observa en la Tabla 4 que la visual es la dominante para todas las emociones a excepción de la expresión emocional neutra, donde el reconocimiento de la prosodia de mensajes acústicos adquiere igualmente un papel importante. Tres de los 26 tests con estímulos presentados por vía visual exponen emociones en secuencias dinámicas. El resto presenta estímulos estáticos (fotografías de expresiones faciales). De forma exclusiva, un test – *el Fantie's Cartoon test* de Fantie (1989)— recoge el contexto donde se desarrolla la emoción que debe seleccionarse. Solo dos tests miden reconocimiento emocional visual por medio de cambios en la expresión de la mirada. Estos son el *The mind in the eyes test* de Baron-Cohen, Wheelwright, Hill, Raste y Plumb (2001), que efectúa mediciones sólo con presentaciones de miradas, y el *The Frankfurt Test and Training of Facial Affect Recognition, FE-FA* (Bölte, Feineis-Matthews y Poustka, 2003 citado en Kuusikko et al. 2009) que utiliza también caras completas.

De los 33 instrumentos de medida hallados, no hay ninguno cuyas puntuaciones se hayan analizado con el Modelo de Rasch y solo uno se analiza mediante el modelo de dos parámetros, el desarrollado por Suzuki et al. (2006). No se encuentra información psicométrica referente a 11 de los 33 tests revisados. En seis de los instrumentos se utiliza más de un procedimiento de estimación de la fiabilidad, aunque lo habitual es informar de la consistencia interna: nueve tests aportan índices de  $\alpha$  de Cronbach con valores aceptables, por encima de .70; siete de los instrumentos miden la fiabilidad por medio de medidas test-retest; otros siete tests analizan también el acuerdo inter-observadores. En todos los casos el valor de los índices es adecuado (por encima de .70).

El análisis de la validez de los instrumentos se realiza en un caso –el test *The Multimodal Emotion Recognition Test (MERT)* de Bänziger et al. (2009) – por medio de un análisis factorial exploratorio de las puntuaciones para poner a prueba la unidimensionalidad de la medida. El test MERT se compone de una parte de reconocimiento emocional por vía visual y otra parte por vía acústica. Los autores confirman en el análisis factorial que el reconocimiento emocional facial y acústico son dos habilidades distintas e independientes entre sí. Doce tests estiman la validez de la medida a partir del análisis de la correlación de las puntuaciones obtenidas por los sujetos con otras puntuaciones en otras escalas similares de reconocimiento emocional (validez concurrente). En todos los casos, existe evidencia de validez de constructo. Tres tests aportan datos neuropsicológicos para corroborar la validez de sus puntuaciones. Esa evidencia procede de la práctica clínica pues estos tres instrumentos tuvieron suficiente sensibilidad en evaluaciones previas como para hacer distinciones entre muestras normales y muestras con daño cerebral. A su vez, permiten establecer diferencias en el rendimiento de pacientes con daño cerebral en el hemisferio izquierdo y derecho. Esos tests son el *Fantie's Cartoon test* de Fantie (1989), el *Kinney's Affect Matching Test (KAMT)* de Kinney, Fantie, Pascualvaca, Rodríguez y Mirsky (1995) y el *Affective Prosody Test (APT)* de Hiou, Vagia, Haritidou, Karakostas y Kosmidis (2004).

Con respecto a la sensibilidad de los instrumentos de medida para detectar diferencias significativas entre las distintas muestras en la dimensión analizada, 26 tests lo hacen satisfactoriamente mientras que en cinco de ellos no se encuentran diferencias entre las muestras que responden a la prueba. Estos cinco tests son el *Affective prosody test* (Hiou, Vagia, Haritidou, Karakostas y Kosmidis, 2004, citado en Bozikas, 2009), el *Fantie's Cartoon test* de Fantie (1989), el *Kinney's Affect Matching Test (KAMT)* de Kinney, Fantie, Pascualvaca, Rodríguez, y Mirsky (1995), un test informatizado desarrollado por Murray (1999, citado en Venn, Gray, Montagne, Murray, Burt, Frigerio, Perrett y Young, 2004) y el *Diagnostic Analysis of Nonverbal Accuracy (DANVA-2)* desarrollado por Nowicki y Carton (1993). Los tres primeros se usan en el mismo estudio (Bozikas et al., 2009) para establecer distinciones entre una muestra con trastorno obsesivo compulsivo y un grupo control. El cuarto test se emplea en una investigación que compara a una muestra con trastorno bipolar y a un grupo control en reconocimiento emocional. El quinto test (*DANVA-2*) es una versión del *DANVA* original desarrollado por Nowicki y Duke (1994) y no logra hacer distinciones en reconocimiento emocional entre pacientes con traumatismo craneoencefálico y un grupo control en la parte que contiene estímulos (caras y voces) infantiles, no con las adultas. No se cuenta con información sobre la sensibilidad de dos instrumentos de medida – el test *JACBART* (Matsumoto et al., 2000) y el *Facial Discrimination Task* (Rojahn et al., 2000) al centrarse toda la descripción en la fiabilidad y validez de los instrumentos y no en su sensibilidad.

## Conclusiones

El objetivo principal de esta revisión sistemática era efectuar un análisis de los instrumentos dedicados a la medida del reconocimiento y la expresión de emociones. El proceso de selección y análisis de los distintos estudios se realizó en dos fases, desarrollando sistemas de categorías *ad hoc* previamente a la fase de cuantificación.

Los resultados de la primera fase mostraron la escasa cantidad de artículos centrados en el tema del reconocimiento y, especialmente, la expresión emocional. De los 721 resúmenes seleccionados una vez eliminados los solapamientos existentes entre las bases de datos de Scopus y PsycInfo, sólo 166 versan sobre el tema específico del reconocimiento y/o la expresión emocional. Únicamente 36 de esos artículos mencionan medidas para alguna de las dos variables analizadas. Dentro de los 166 estudios de reconocimiento emocional, hay 41 que emplean medidas no estandarizadas de reconocimiento y/o expresión emocional en temas de relevancia para la Neuropsicología Clínica como el daño cerebral, la demencia (frontotemporal o tipo Alzheimer) o la epilepsia temporal.

Dado que sólo hay un test de expresión emocional dentro de los 33 tests analizados en la segunda fase de la revisión, los resultados de aquí en adelante se discuten básicamente para el reconocimiento. Los resultados revelan patrones de datos referidos a la naturaleza de las muestras, el tipo de estímulos empleados y los procesos de evaluación. Ha quedado patente la relevancia de los tests de reconocimiento con muestras patológicas de edad adulta. Los estudios realizados con la finalidad de realizar revisiones psicométricas de las medidas suelen llevarse a cabo con muestras normales mientras que aquellos con objetivos correlacionales emplean muestras patológicas. Es destacable la cantidad de artículos –trece de los 36 estudios analizados– que utilizan en algún momento los estímulos o los instrumentos de medida creados por Ekman y Friesen (1976) o Matsumoto y Ekman (1989). Si ya es escaso el volumen de investigación centrado en la puesta a prueba de tests de comunicación no verbal de tipo afectivo, más lo es aún la cantidad de estudios centrados en la creación o al menos en el empleo de instrumentos de medida distintos a los clásicos ya mencionados. Tal y como afirman Schaffer, Wisniewski, Dahdah y Froming (2009), a pesar de que haya algunos intentos de desarrollar medidas de reconocimiento emocional, parece que solo unos pocos de esos instrumentos consiguen popularizarse en Neuropsicología.

Con respecto al tipo de evaluación realizado en cada instrumento, resultó notable la proporción de tests que introducen variaciones en el tipo de emoción y que miden, a su vez, el reconocimiento de tipos de emoción por parte del examinando. A su vez, los estímulos (caras o voces) se presentan prácticamente siempre fuera de un contexto que sirva de marco a la emoción expuesta. Las emociones más utilizadas en los instrumentos de medida son la alegría, la pena y la rabia en comparación a otras casi inexistentes en los tests,

como el desprecio. Teniendo en cuenta que Russell afirmó en el año 1994 que la alegría es la emoción más fácil de identificar, es llamativa su omnipresencia en los tests de reconocimiento emocional cuando, probablemente, no ayude a establecer diferencias entre sujetos con una capacidad media-alta de reconocimiento emocional (efecto techo). La alegría es la única emoción que, según Suzuki et al. (2006), escapa al efecto de las medidas propuestas por Matsumoto et al. (2000) para mitigar el efecto techo en los tests de reconocimiento emocional. Estas medidas suponen, por ejemplo, reducir el tamaño de la imagen de la expresión facial, distorsionar su calidad o aumentar la velocidad de presentación. Además del excesivo empleo de la alegría en las medidas de procesamiento emocional, es llamativo el escaso uso que se hace del desprecio en los estudios teniendo en cuenta la relevancia de esta emoción en Psicología (Delgado, 2009; Haidt, 2007).

Más preocupante aún que la presencia de limitaciones metodológicas como la anterior en el diseño de tests, es el escaso uso de los nuevos modelos de medida que detectarían

su existencia. Por ejemplo, el Modelo de Rasch permitiría detectar problemas como la falta de unidimensionalidad esencial o el funcionamiento diferencial de ciertos ítems en relación con variables de la muestra tales como el estatus clínico o la edad. Con estos indicadores de la Psicometría clásica, desconocemos las propiedades de los ítems que componen cada instrumento y la utilidad de cada uno para discriminar entre las muestras de sujetos. Teniendo en cuenta la gravedad de esta limitación, es el momento de proponer la implementación de los nuevos modelos de medición psicométrica en Neuropsicología y, más específicamente, en una variable con tantas implicaciones a nivel básico y aplicado como el reconocimiento de emociones.

Resulta también llamativo que sólo tres de los 33 tests analizados aporten datos neuropsicológicos para dar cuenta de la validez de sus mediciones. Un claro avance desde el punto de vista científico sería considerar en mayor medida los datos neuropsicológicos al construir o interpretar los resultados de tests de reconocimiento emocional.

## Referencias

- Bänzinger, T., Grandjean, D. y Sherer, K. R. (2009). Emotion recognition from expressions in face, voice and body: The Multimodal Emotion recognition test (MERT). *Emotion*, 8, 691-704.
- Baron-Cohen, S., Wheelwright, S., Spong, A., Scahill, V. y Lawson, J. (2001). Are intuitive physics and intuitive psychology independent? A test with children with Asperger Syndrome. *Journal of Developmental and Learning Disorders*, 5, 47-48.
- Bozikas, V. P., Kosmidis, M. H., Giannakou, M., Saitis, M., Fokas, K. y Garyfallos, G. (2009). *Journal of the International Neuropsychological Society*, 15, 148-153.
- Cadavid, N., Delgado, A.R. y Prieto, G. (2007). Construcción de una escala de depresión con el modelo de Rasch. *Psicothema*, 19, 515-521.
- Darwin, Ch. (1872). *The Expression of the Emotions in Man and Animals*. London: John Murray.
- Delgado, A.R. (2009). Spanish Basic Emotion Words are Consistently Ordered. *Quality & Quantity*, 43, 509-517.
- Ekman, P. y Friesen, W. V. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.
- Erwin, R. J., Gur, R. C., Gur, R. E., Skolnick, B., Mawhinney-Hee, M. y Smalls, J. (1992). Facial emotion discrimination 1: Task construction and behavioral findings in normal subjects. *Psychiatry Research* 1992, 42, 231-240.
- Fantie, B. (1989). *Fantie's Cartoon Test*. Washington, D.C: Human Neuropsychology Laboratory, American University.
- Gur, R. C., Sara, R., Hagendoorn, M., Marom, O., Hughett, P., Macy, L., Turner, T., Bajcsy, R., Posner, A. y Gur, R. E. (2002). A method for obtaining 3 dimensional facial expressions and its standardization for use in neurocognitive studies. *Journal of Neuroscience Methods*, 115, 137-143.
- Haidt, J., (2007). The new synthesis in moral psychology. *Science*, 316, 998-1002.
- Kinney, J. M., Fantie, B. D., Pascualvaca, D. M., Rodriguez, D. y Mirsky, A. F. (1995). Comprehension of affect in children with pervasive developmental disorders: Deficits in matching faces to cartoon situations. *Journal of the International Neuropsychological Society*, 1, 156-157.
- Kuusikko, S., Haapsamo, H., Jansson-Verkasalo, E., Hurtig, T., Mattila, M. L., Ebeling, H., Jussila, K., Bölte, S. y Moilanen, I. (2009). Emotion Recognition in Children and Adolescents with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, 39, 938-945.
- Lopes, P., Delgado, A.R., Prieto, G., Gamito, P. y Trigo, H. (2010). Rasch-modeling the Portuguese SOCRATES in a clinical sample. *Psychology of Addictive Behaviors*, 24, 355-359.
- Mathersul, D., Palmer, D. M., Gur, R. C., Gur, R. E., Cooper, N., Gordon, E. y Williams, L. M. (2009). Explicit identification and implicit recognition of facial emotions: II. Core domains and relationships with general cognition. *Journal of Clinical and Experimental Neuropsychology*, 31, 278-291.
- Matsumoto, D. y Ekman, P. (1988). *Japanese and Caucasian facial expressions of emotion (JACFEE) and neutral faces (JACNeuF)*. San Francisco, CA: San Francisco State University.
- Matsumoto, D., LeRoux, J., Wilson-Cohn, C., Raroque, J., Kooken, K., Ekman, P., Yrizarry, N., Loewinger, S., Uchida, H., Yee, A., Amo, L. y Goh, A. (2000). A new test to measure emotion recognition ability: Matsumoto and Ekman's Japanese and Caucasian Brief Affect Recognition Test (JACBART). *Journal of Nonverbal Behavior*, 24, 179-209.
- Nakhuntina, L., Borod, J. C. y Zgaljardic, D. J. (2006). Posed prosodic emotional expression in unilateral stroke patients: Recovery, lesion location, and emotional perception. *Archives of Clinical Neuropsychology*, 21, 1-13.
- Nowicki, S. y Carton, J. (1993). The measurement of emotional intensity from facial expressions. *Journal of Social Psychology*, 133, 749-750.
- Nowicki, S. y Duke, M. P. (1994). Individual differences in the nonverbal communication of affect. The diagnostic analysis of nonverbal accuracy scale. *Journal of nonverbal behaviour*, 18, 9-35.
- Philips, M. L., Drevets W. C., Rauch S. L. y Lane R. (2003). Neurobiology of emotion perception I: the neural basis of normal emotion perception. *Biological Psychiatry*, 54, 504-514.
- Prieto, G. y Delgado, A.R. (2007). Measuring Math Anxiety (in Spanish) with the Rasch Rating Scale Model. *Journal of Applied Measurement*, 8, 149-160.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: The Danish Institute for Educational Research.
- Rojahn, J., Gerhards, F., Matlock, S. T. y Kroeger, T. L. (2000). Reliability and validity studies of the Facial Discrimination Task for emotion research. *Psychiatry Research* 95, 169-181.
- Schaffer, S. G., Wisniewski, A., Dahdah, M. y Froming, K. B. (2009). The Comprehensive Affect Testing System-Abbreviated: Effects of Age on Performance. *Archives of Clinical Neuropsychology*, 24, 89-104.
- Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the crosscultural studies. *Psychological Bulletin*, 115, 102-141.
- Tonks, J., Williams, W. H., Frampton, I., Yates, P. y Slater, A. (2007). Assessing emotion recognition in 9-15-years olds: Preliminary analysis of abilities in reading emotion from faces, voices and eyes. *Brain Injury*, 21, 623-629.

- Suzuki, A., Hoshino, T. y Shigemasu (2006). Measuring individual differences in sensitivities to basic emotions in faces. *Cognition*, *99*, 327-353.
- Venn, H. R., Gray, J. M., Montagne, B., Murray, L. K., Burt, D. M., Frigerio, E., Perrett, D. I. y Young, A. H. (2004). Perception of facial expressions of emotion in bipolar disorder. *Bipolar Disorders*, *6*, 286-293.
- Williams, L. M., Mathersul, D., Palmer, D. M., Gur, R. C., Gur, R. E. y Gordon, E. (2009). Explicit identification and implicit recognition of facial emotions: I. Age effects in males and females across 10 decades. *Journal of Clinical and Experimental Neuropsychology*, *31*, 257-277.

(Artículo recibido: 25-11-2010, revisión: 21-09-2011, aceptado: 16-10-2011)

### Anexo. Listado de tests con sus correspondientes citas bibliográficas

| Nombre del test  | Fuente bibliográfica  |
|--|---|
| Affective prosody test (APT)   | Hiou, K., Vagia, A., Haritidou, E., Karakostas, D. y Kosmidis, M. H. (2004). Affect perception as a cognitive function: Validity and clinical application of a Neuropsychological Test Battery in healthy individuals and patients with brain lesions. <i>Psychology</i> , <i>11</i> , 388-401.   |
| Carolina older adults test of nonverbal Communications (COAT-NC)       | Spell, L. A. y Frank, E. M. (2000). Recognition of nonverbal communication of affect following traumatic brain injury. <i>Journal of Nonverbal Behavior</i> , <i>24</i> , 285-300.  |
| Comprehensive affect testing system                                    | Froming, K.B, Levy, C. M., Shaffer, S. G. y Ekman, P. (2006). Comprehensive Affect Testing System. <a href="http://www.psychologysoftware.com/testing_instruments.htm">http://www.psychologysoftware.com/testing_instruments.htm</a> .  |
| Diagnostic analysis of nonverbal accuracy (DANVA)                      | Nowicki, S. y Duke, M. P. (1994). Individual differences in the nonverbal communication of affect. The diagnostic analysis of nonverbal accuracy scale. <i>Journal of Nonverbal Behaviour</i> , <i>18</i> , 9-35.   |
| Diagnostic analysis of nonverbal accuracy-2 (DANVA-2)                  | Nowicki, S. y Carton, J. (1993). The measurement of emotional intensity from facial expressions; the DANVA FACES 2. <i>Journal of Social Psychology</i> , <i>133</i> , 749-751.   |
| The Penn emotion differentiation test (Emodiff)                        | Kohler, C. G., Bilker, W., Hagendoorn, M., Gur, R. E. & Gur, R. C. (2000). Emotion-recognition deficit in schizophrenia: Association with symptomatology and cognition. <i>Biological Psychiatry</i> , <i>48</i> , 127-136.   |
| The emotion recognition test (ER40)                                    | Gur, R. C., Schroeder, L., Turner, T., McGrath, C., Chan, R. M., Turetsky, B.I., Alsop, D., Maldjian, J. y Gur, R. E. (2002). Brain activation during facial emotion processing. <i>NeuroImage</i> , <i>16</i> , 651-662.   |
| Emotion recognition test (ERT)   | Shimokawa, A., Yatomi, N., Anamizu, S., Torii, S., Isono, H., Sugai, Y. y Kohno, M. (2001). Influence of deteriorating ability of emotional comprehension on interpersonal behavior in Alzheimer-type dementia. <i>Brain and Cognition</i> , <i>47</i> , 423-433.   |
| The expression recognition hexagon stimuli                             | Calder, A. J., Young, A. W., Rowland, D. y Perrett, D. I. (1996). Facial emotion recognition after bilateral amygdala damage: differentially severe impairment of fear. <i>Cognitive Neuropsychology</i> , <i>13</i> , 699-745.   |
| Recognition of facial affect (FACES)                                   | Haskins, B., M. Shutty, and E. Kellogg. (1995). Affect Processing in Chronically Psychiatric Patients: Development of a reliable assessment tool. <i>Schizophrenia Research</i> , <i>15</i> , 291-297.  |
| Facial discrimination task (FDT)                                       | Erwin, R. J., Gur, R. C., Gur, R. E., Skolnick, B., Mawhinney-Hee, M. y Smalls, J. (1992). Facial emotion discrimination: 1. Task construction and behavioral findings in normal subjects. <i>Psychiatry Research</i> , <i>42</i> , 231-240.  |
| Facial emotion recognition task  | Chan, C.H., Wong, R., Wang, K. y Lee, T.M.C. (2008). Emotion recognition in chinese people with schizophrenia. <i>Psychiatry Research</i> , <i>157</i> , 67-76.   |
| Fantie's Cartoon Test.   | Fantie, B. (1989). <i>Fantie's Cartoon Test</i> . Washington, D.C: Human Neuropsychology Laboratory, American University  |
| Facially expressed emotion labeling (FEEL-Test)                        | Kessler, H. (2001). <i>Entwicklung und Reliabilitätsstudie des FEEL-Tests (Facially Expressed Emotion Labeling)</i> . (Unpublished doctoral dissertation.) Universität Ulm.   |
| Frankfurter test und training des erkennens von fazialem affekt (FEFA) | Bölte, S., Feineis-Matthews, S. y Poustka, F. (2003). <i>Frankfurter test und training des erkennens von fazialem affekt, FEFA</i> . Frankfurt am Main: J.W. Goethe-Universitätsklinikum.   |
| Facial emotion perception task (FEPT)                                  | Rapport, L. J., Friedman, S. L., Tzelepis, A. y VanVoorhis, A. (2002). Experienced emotion and affect regulation in adult attention-deficit hyperactivity disorder. <i>Neuropsychology</i> , <i>16</i> , 102-110.   |
| Florida affect battery   | Bowers, D., Blonder, LX, & Heilman, KM (1999). <i>Florida Affect Battery</i> . Gainesville: Center for Neuropsychological Studies, University of Florida.   |
| Interactive computer programme   | Murray, L. M. (1999). <i>Facial expression perception in different psychiatric groups</i> . (Unpublished doctoral dissertation). University of St Andrews.  |
| Japanese and Caucasian brief affect recognition test (JACBART)         | Matsumoto, D., LeRoux, J., Wilson-Cohn, C., Raroque, J., Kookan, K., Ekman, P., Yri-zarry, N., Loewinger, S., Uchida, H., Yee, A., Amo, L. y Goh, A. (2000). A new test to measure emotion recognition ability: Matsumoto and Ekman's Japanese and Caucasian Brief Affect Recognition Test (JACBART). <i>Journal of Nonverbal Behavior</i> , <i>24</i> , 179-209. |
| Kinney's Affect Matching Test (KAMT)                                   | Kinney, J. M., Fantie, B. D., Pascualvaca, D. M., Rodriguez, D. y Mirsky, A. F. (1995). Comprehension of affect in children with pervasive developmental disorders: Deficits in matching faces to cartoon situations. <i>Journal of the International Neuropsychological Society</i> , <i>1</i> , 156-157.  |

|  |   |
|--|---|
| The Multimodal Emotion recognition test (MERT) | Bänzinger, T., Grandjean, D. y Sherer, K. R. (2009). Emotion recognition from expressions in face, voice and body: The Multimodal Emotion recognition test (MERT). <i>Emotion</i> , 8, 691-704.   |
| The NimStim set of facial expressions          | Marcus, D. J., Westerlund, A., Casey, B. J, Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. <i>Psychiatry Research</i> , 168, 242-249. <a href="http://www.macbrain.org/resources.htm">http://www.macbrain.org/resources.htm</a>    |
| The New York Emotion Battery                   | Borod, J. C., Welkowitz, J. y Obler, L. K. (1992). <i>The New York Emotion Battery</i> . Unpublished materials, Mount Sinai Medical Center, Department of Neurology, New York.  |
| The Penn emotion acuity test (PEAT)            | Kohler, C. G., Bilker, W., Hagendoorn, M., Gur, R. E. & Gur, R. C. (2000). Emotion-recognition deficit in schizophrenia: Association with symptomatology and cognition. <i>Biological Psychiatry</i> , 48, 127-136.   |
| Perception of emotion test (POET)              | Egan, G. J. (1989). <i>Assessment of emotional processing in right and left hemisphere stroke patients: A validation study of the Perception of Emotion Test</i> . (Unpublished doctoral dissertation). Georgia State University.   |
| Pictures of facial affect                      | Ekman, P. y Friesen, W. V. (1976). <i>Pictures of facial affect</i> . Palo Alto, CA: Consulting Psychologists Press.  |
| Prosodic recognition task                      | Chan, C.H., Wong, R., Wang, K. y Lee, T.M.C. (2008). Emotion recognition in chinese people with schizophrenia. <i>Psychiatry Research</i> , 157, 67-76.   |
| Prosodic comprehension (PROSCOMP)              | Haskins, B., M. Shutty, and E. Kellogg. (1995). Affect Processing in Chronically Psychiatric Patients: Development of a reliable assessment tool. <i>Schizophrenia Research</i> , 15, 291-297.  |
| The "Reading the mind in the eyes" test        | Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y. y Plumb, I. (2001). The "Reading the mind in the eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. <i>Journal of Child Psychiatry and Psychology</i> , 42, 241-252. |
| Sensitivities to basic emotions in faces       | Suzuki, A., Hoshino, T. y Shigemasa (2006). Measuring individual differences in sensitivities to basic emotions in faces. <i>Cognition</i> , 99, 327-353  |
| Social-emotional questionnaire (SEQ)           | Bramham, J., Morris, R. G., Hornak, J., Bullock, P. y Polkey, C. E. (2009). Social and emotional functioning following bilateral and unilateral neurosurgical prefrontal cortex lesions. <i>Journal of Neuropsychology</i> , 3, 125-143.  |
| The awareness of social inference test (TASIT) | McDonald, S., Flanagan, S., y Rollins, J. (2001). <i>The Awareness of Social Inference Test</i> . Sydney, Australia: Harcourt Assessment.   |
| Webneuro (Tests of emotion processing)         | Silverstein, S. M., Berten, S., Olson, P., Paul, R., Williams, L.M., Cooper, N. y Gordon, E. (2007). Development and validation of a World-Wide-Web-based neurocognitive assessment battery: WebNeuro. <i>Behaviour Research Methods</i> , 39, 940-949.                                       |