

Elena Huerta-Ramos^{a,b,c,d,*},
Marta Ferrer-Quintero^{a,b,c,g,*},
Juana Gómez-Benito^h,
Fermín González-Higueras^g,
Daniel Cuadras^c,
Ángel Luis del Rey-Mejías^{d,e,f},
Judith Usall^{a,b,c},
Susana Ochoa^{a,b,c,d}

TRADUCCIÓN Y VALIDACIÓN DEL TEST DE CARAS DE BARON COHEN EN POBLACIÓN ESPAÑOLA

**Las dos autoras realizaron la misma contribución a la preparación y redacción del presente manuscrito.*

- a) Institut de Recerca de Sant Joan de Déu. Edifici Docent. Santa Rosa, 39-57 08950 Esplugues del Llobregat (Barcelona). Spain.
- b) Parc Sanitari Sant Joan de Déu, Sant Boi de Llobregat, Spain. Carrer del Doctor Antoni Pujadas, 42, 08830 Sant Boi de Llobregat.
- c) Fundació Sant Joan de Déu, Esplugues de Llobregat, Spain. Edifici Docent. Santa Rosa, 39-57 08950 Esplugues del Llobregat (Barcelona). Spain.
- d) CIBERSAM (Centro de Investigación Biomédica en Red - Salud Mental). Av. Monforte de Lemos, 3-5. Pabellón 11. Planta 0 28029 Madrid (Spain).
- e) Hospital Clínico San Carlos, Madrid, Spain. C/ Profesor Martín Lagos, s/n. 28040 Madrid (Spain).
- f) Comunidad Terapéutica de Jaén. Avda. de Madrid, s/n, Jaén, 23008 (Spain)
- g) Departamento de Metodología, Facultad de Psicología, Universidad Complutense de Madrid. Campus de Somosaguas, 28223, Pozuelo de Alarcón. Madrid, Spain
- h) Departament de Psicologia Social i Psicologia Quantitativa, Universitat de Barcelona. Passeig de la Vall d'Hebron, 171, 08035 Barcelona, Spain.

Elena Huerta-Ramos diseñó el proyecto. Marta Ferrer-Quintero, Fermín González-Higueras y Ángel Luis del Rey-Mejías recolectaron los datos. Marta Ferrer-Quintero y Elena Huerta-Ramos escribieron el manuscrito. Daniel Cuadras realizó el análisis estadístico. Susana Ochoa, Judith Usall y Juana Gómez-Benito supervisaron las partes teóricas, metodológicas y estadísticas, además de ayudar con la interpretación de los resultados.

RESUMEN

Introducción. El reconocimiento facial de emociones es esencial en el funcionamiento social adecuado. Se han encontrado déficits en muchas poblaciones clínicas. Hay pocos tests validados que midan esta habilidad y ninguno en población española. En este trabajo, traducimos y validamos el Test de Caras de Baron Cohen en población general española.

Métodos. El test fue administrado a 211 personas (63,3 % mujeres) sanas de entre 19 y 70 años de edad. Usamos matrices tetracóricas para obtener la fiabilidad test-retest y la consistencia interna. Se realizó un análisis factorial confirmatorio para comprobar la unidimensionalidad del test. Utilizamos correlaciones de Pearson para examinar asociaciones entre variables.

Resultados. La media en el estudio fue de 18 (DE = 1,38). Se obtuvo un alfa de Cronbach de 0,75. Calculamos los índices Guttman Lambda 3 para cada ítem. 17 de 20 ítems obtuvieron una estabilidad test-retest excelente. No encontramos asociaciones entre el rendimiento y el género, la edad o el nivel académico. El test presentó una estructura unidimensional (CFI = 0,889; TLI = 0,873 y RMSEA = 0,047).

Conclusiones. El Test de Caras de Baron Cohen puede ser útil como instrumento de medida a pesar de no ser sensible al género y a la edad. Puesto que presenta un efecto techo, no resulta un instrumento adecuado para obtener medidas precisas del funcionamiento superior de esta habilidad.

Palabras clave. Test de Caras, reconocimiento facial de emociones, validación, propiedades psicométricas, población general.

Autoras de correspondencia:

Elena Huerta-Ramos: mehuerta@pssjd.org
Marta Ferrer Quintero: mf.quintero@pssjd.org

Actas Esp Psiquiatr 2021;49(3):106-13 | ISSN: 1578-2735

TRANSLATION AND VALIDATION OF BARON COHEN'S FACE TEST IN A GENERAL POPULATION FROM SPAIN

ABSTRACT

Introduction. Facial emotion recognition is considered the foundation of effective social functioning, but it has been found impaired in several clinical populations. However, there are few validated tests to measure the ability. To the best of our knowledge, there is no validated measure in a Spanish population. We translated and validated Baron Cohen's Face Test in a general Spanish population.

Methods. The test was administered to 211 (63.3% female) healthy volunteers between 19 and 70 years of age. We used tetrachoric matrices to obtain item per item test-retest reliability and internal consistency. We used confirmatory factor analysis to test for unidimensionality. We used Pearson correlations to examine associations between variables.

Results. The mean score was 18 (SD=1.38). Cronbach's alpha was 0.75. Guttman Lambda 3 indexes yielded 17 out of 20 items to have excellent test-retest reliability. Gender or age differences in performance were not found. The test seems to comply with a one-dimensional structure: CFI=0.889; TLI=0.873 and RMSEA=0.047.

Conclusions. Baron Cohen's Face Test could be a valid measure of FER, although it is not sensitive to age or gender. Because it presents a certain ceiling effect, it could not be appropriate to detect excellent performance.

INTRODUCCIÓN

El reconocimiento facial de emociones (FER) es crucial en el buen funcionamiento social. Se ha explorado exhaustivamente tanto en población sana¹⁻³ como en diferentes poblaciones clínicas⁴⁻⁷.

A pesar del numeroso volumen de publicaciones sobre este tema, muchos de los instrumentos empleados en su medición tienen propiedades psicométricas pobres o incluso desconocidas. Se han validado algunas bases de datos de fotografías (ver⁸⁻¹⁴ para diferentes ejemplos) que pueden ofrecer gran control experimental en diferentes variables como la edad, el género, la etnia y la validez ecológica. Sin embargo, son demasiado grandes y su uso ha conllevado diferencias metodológicas que han impedido la comparación de los resultados.

Las propiedades psicométricas de la mayoría de las tareas que evalúan el FER son pobres¹⁵ y carecen de validaciones en diferentes poblaciones. Algunas de ellas son el "Ekman-60 Faces Test" (Dodich *et al.*, 2014), el "Japanese and Caucasian Brief Affective Recognition Test" (JACBARTT)¹⁷, el "Facial Emotion Recognition Test" (FERT)¹⁸, el "Bell Lysaker Emotion Recognition Test" (BLERT)¹⁹, el "Reading the Mind in the Eyes Test" (RMET)²⁰ o el "The Videotest of Emotion Recognition"²¹. Puesto que realizar una revisión exhaustiva de las tareas y tests que evalúan el FER está más allá de los objetivos del presente trabajo, sugerimos consultar Passarelli *et al.*, 2018 para una revisión completa.

Algunas de estas tareas han sido validadas en algunos países pero no en otros. Algunas han sido ampliamente utilizadas en investigación sin haberse realizado un trabajo de validación previo. Además, muchas de estas tareas son excesivamente largas o exhaustivas en según qué contextos. Una versión reducida del "The Assessment of Social Inference Test" (TASIT)²² ha demostrado su fiabilidad como medida de cribado²³, pero existe poca literatura explorando pruebas breves, pero robustas y sensibles, que evalúen otros dominios dentro de la cognición social.

En población sana, ha quedado demostrado que las mujeres presentan mejor reconocimiento facial de emociones tanto en la infancia como en la adolescencia y la edad adulta^{24,25}. A pesar de algunos resultados poco consistentes, los resultados de un metaanálisis sugieren que el FER decae con la edad, manteniéndose intacta la habilidad para reconocer el asco, resultados que son consistentes con el envejecimiento normal del cerebro^{26,27}.

La influencia del nivel académico en FER se ha estudiado menos que otras variables demográficas. El único estudio que específicamente exploró una asociación encontró una

correlación positiva entre el nivel académico y los resultados en el test FERT y una interacción significativa entre la edad y la escolarización, favoreciendo a las personas más jóvenes y con más años de educación formal²⁸.

Hasta donde sabemos, hay muy pocos tests de FER validados en población española.

Recientemente se ha validado el RMET^{29,30}, pero sus propiedades psicométricas resultaron ser cuestionables. Otro test validado en población española es el "Mayer-Salovey-Caruso Emotional Intelligence Test" (MSCEIT)³¹, aunque debe notarse que el MSCEIT evalúa todos los dominios de la cognición social y no únicamente FER.

El Test de Caras de Baron Cohen (Test de Caras)²⁰ es un test de FER que se desarrolló durante la construcción del RMET. Este test es de libre uso en contextos de investigación y clínicos y puede ser descargado de Autism Research Centre (https://www.autismresearchcentre.com/arc_tests).

De acuerdo con dicha página web, este test se ha traducido a más de diez idiomas, pero a pesar de nuestras búsquedas, no hemos podido encontrar ningún artículo de validación o de valores de referencia, ni en poblaciones sanas ni en poblaciones clínicas. Es posible que este test esté siendo utilizado por personal clínico e investigador desconociendo sus propiedades psicométricas. Además, se sabe que las diferencias culturales y lingüísticas pueden influenciar medidas neuropsicológicas³², por lo que deben tenerse en cuenta estas diferencias a la hora de interpretar el resultado.

De acuerdo a lo expuesto, nuestro objetivo con este estudio es validar el Test de Caras de Baron Cohen en la población española, obtener las propiedades psicométricas del instrumento y explorar su sensibilidad a variables demográficas.

MÉTODOS

El Test de Caras fue traducido y adaptado de acuerdo a las guías de la Comisión Internacional del Test (ITC)³³.

Dado que el test no incluye ninguna frase completa y se compone únicamente de palabras, un equipo de tres personas (bilingüe español-inglés, nativo español y nativo inglés) tradujo las palabras de acuerdo a su frecuencia de uso, formas europeas del lenguaje y género del estímulo. Se obtuvo una versión piloto del test, que fue revisada por personal investigador para la detección de errores y su posterior corrección.

Los criterios de inclusión de los participantes incluyeron: tener una edad entre 18 y 70 años y firmar el consentimiento

to informado electrónico. Los criterios de exclusión contemplaron: enfermedad mental en el momento del estudio, enfermedad mental grave, discapacidad intelectual, trastornos del desarrollo, daño neurológico o demencia.

La muestra fue reclutada mediante el método bola de nieve por 3 miembros del estudio en 3 hospitales pertenecientes a diferentes regiones españolas: Parc Sanitari Sant Joan de Déu, Comunidad Terapéutica de Jaén y Hospital Clínico San Carlos. Se utilizó este método para asegurar la mayor participación de personas de la comunidad de diferentes regiones de España. Los primeros participantes fueron escogidos del círculo social inmediato del equipo investigador. Recibieron una encuesta online que podía ser completada desde cualquier tipo de aparato electrónico. La encuesta se compuso de los tests descritos abajo, un cuestionario socio-demográfico y un cuestionario sobre antecedentes de salud mental. Se animó a los primeros participantes a diseminar el cuestionario por su entorno social. Las respuestas fueron inspeccionadas individualmente. Las personas que en el momento de la encuesta recibían medicación psicoactiva fueron excluidas, con la excepción del uso de benzodiazepinas como relajante muscular.

La muestra fue reclutada entre julio de 2016 y enero de 2017 (incluyendo re-test). Tres meses tras la primera administración, se pidió a 37 participantes que completaran una encuesta re-test, que incluía el Test de Caras y el RMET. Los participantes fueron seleccionados a través de una lista aleatoria. De los 37 participantes, 24 contestaron a la encuesta re-test. Los datos de todos los participantes fueron alojados en un servidor del hospital Parc Sanitari Sant Joan de Déu

que cumple con todos los requerimientos de seguridad para datos de salud e investigación.

Test de Caras de Baron Cohen³⁴: el test consiste en 20 ítems que muestran la fotografía de una actriz expresando una emoción. Los participantes deben escoger qué emoción está sintiendo entre dos opciones de respuesta. La mitad de los ítems consisten en emociones básicas, mientras que los otros 10 ítems son estados mentales complejos. Ejemplos de cada estímulo se muestran en la figura 1.

El "Reading the Mind in the Eyes Test"³⁴: Consiste en 36 imágenes circunscritas a la región de los ojos. Los participantes deben elegir la emoción que mejor está expresando los ojos entre 4 diferentes opciones de respuesta. Se ofrece un glosario de cada término en el caso de que la persona participante desconozca el significado de alguna de las palabras.

Análisis de datos

Los datos fueron analizados con IBM SPSS Statistics 22 y R³⁵. Se utilizaron matrices tetracóricas para calcular el alfa de Cronbach, la estabilidad test-retest y el análisis factorial. Se utilizó esta técnica para asegurar un mejor ajuste al utilizar ítems binarios. Se utilizó el alfa de Cronbach como medida de la consistencia interna. Se utilizó Guttman Lambda 3 para obtener la estabilidad temporal de cada ítem, puesto que este método es más adecuado en ítems binarios que el coeficiente de correlación intraclass. Se utilizó una correlación de Pearson entre el total de cada aplicación para obtener la fiabilidad test-retest del test completo. La validez convergente se examinó con



SORPRENDIDA



CONTENTA



ENFADADA



ASUSTADA

Figura 1

Estímulo 3, "sorprendida" vs. "contenta" y 4, "enfadada" vs. "asustada"

una correlación de Pearson entre el total del Test de Caras y el total del RMET. Se utilizó el t de Student para medidas independientes para calcular las diferencias entre variables. Se utilizó ANOVA para encontrar diferencias entre grupos. Se usaron correlaciones de Pearson para encontrar correlaciones entre variables. Se asumió un intervalo de confianza del 95 %.

Requerimientos éticos

El estudio se diseñó de acuerdo a la Declaración de Helsinki de la Asociación Médica Mundial³⁶ y fue aprobado por el comité de ética del hospital Parc Sanitari Sant Joan de Déu como centro coordinador.

RESULTADOS

Un total de 286 personas recibieron la encuesta. De ellos, 8 sujetos fueron excluidos debido a problemas de salud mental y 67 comenzaron a responder, pero no completaron la encuesta, por lo que sus datos no pudieron ser analizados. La muestra final incluyó 211 participantes, 134 mujeres y 77 hombres con edades comprendidas entre 19 y 70 años.

| Tabla 1 | Características sociodemográficas de la muestra |
|----------------------|---|
| Edad (Media, DE) | 40,01 (13,07) |
| Género (%) | |
| Mujer | 63,3 |
| Hombre | 36,7 |
| Nivel académico (%) | |
| Primaria | 5,8 |
| Secundaria | 16,4 |
| Diplomatura | 13 |
| Licenciatura o grado | 39,1 |
| Máster o posgrado | 20,3 |
| Doctorado | 5,3 |

Los participantes fueron reclutados desde 15 comunidades autónomas del Estado español.

Datos normativos y propiedades psicométricas del Test de Caras

Nuestra muestra obtuvo una puntuación media de 18 (DE = 1,38), distribuidas entre una puntuación mínima de 15 y una máxima de 20.

Para calcular la consistencia interna, se calculó el alfa de Cronbach a partir de la tabla de matrices tetracóricas. Se obtuvo un valor de 0,75, lo que indica una buena consistencia interna. Este valor no podría aumentarse significativamente eliminando ningún ítem. Se calculó la validez convergente obteniendo la correlación entre el total del Test de Caras y del RMET, con la que se obtuvo una correlación baja pero significativa ($r = 0,192, p < 0,005$).

La estabilidad test-retest se calculó obteniendo la correlación entre el total en el test en la primera aplicación y el total en el test en la segunda aplicación ($r = 0,372, p = 0,088$). Posteriormente, se obtuvo la matriz de correlaciones tetracóricas para calcular el índice Guttman Lambda 3 para cada ítem. Este índice es equivalente al alfa de Cronbach³⁷. Los índices de concordancia para cada ítem están representados en la tabla 2.

| Tabla 2 | Índices de acuerdo para la estabilidad test-retest |
|--------------------------|--|
| Índices Guttman Lambda 3 | Ítems |
| > 0,8 | 1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 13, 14, 15, 16 |
| 0,60 - 0,8 | 7, 17, 19 |
| 0,40 - 0,60 | 12, 20 |
| 0,20 - 0,40 | 18 |

Dimensionalidad del instrumento

Se realizó un análisis factorial confirmatorio (CFA) para todos los ítems. No se incluyó el ítem 1, ya que tenía varianza negativa, ni los ítems 3 y 4, ya que eran constantes. Únicamente, se probó la solución de un factor, ya que este es el modelo propuesto por los autores del test. Se utilizó el estimador de mínimos cuadrados ponderados diagonalmente (DWLS), para el que se obtuvo un índice de ajuste comparativo (CFI) de 0,889, un índice de Tucker-Lewis (TLI) de 0,873 y un RMSEA (Root Mean Square Error of Approximation) de 0,047.

Variables moderadoras: género, edad y nivel educativo

No se encontraron diferencias entre géneros ($t(209) = 0,865; p = 0,388$). Las correlaciones entre la puntuación total del test y la edad no fueron significativas ($r = -0,075, p = 0,280$). Se realizó un ANOVA de una vía entre el nivel educativo y la puntuación total en el Test de Caras. No se encontraron diferencias significativas entre los grupos ($F(5,200) =$

1,886; $p = 0,098$), pero se encontró una asociación significativa entre la puntuación total del test y el nivel académico ($r = 0,153$; $p = 0,028$).

DISCUSIÓN

Este trabajo presenta la validación del Test de Caras junto con sus valores normativos en población sana española. Se han obtenido las propiedades psicométricas del test y se discuten las variables moderadoras.

Los valores normativos en el test coinciden con resultados previamente encontrados en población general, que reportan que el reconocimiento facial de emociones funciona bien en población general y se obtienen puntuaciones muy altas en los tests con desviaciones estándar relativamente bajas^{26,38}.

Nuestra muestra obtuvo puntuaciones similares a las halladas en el test original³⁴, que encontraron una media de 9,13 (DE = 0,96) en emociones básicas y una media de 9,38 (DE = 0,62) en estados mentales complejos. Se encontró un efecto techo en nuestra muestra, que puede deberse a que el test es corto y solo existen dos opciones de respuesta para cada ítem, lo que puede haber facilitado que algunos participantes hayan acertado el ítem por azar incluso si su reconocimiento facial de emociones se encuentra comprometido.

En lo referente a sus propiedades psicométricas, se obtuvo una consistencia interna de $\alpha = 0,75$. Otros instrumentos similares han obtenido valores similares: el JACBART tiene una consistencia interna entre 0,86 y 0,92¹⁷, el "Videotest of Emotion Recognition" reportó dos alfas de Cronbach, uno para el índice de precisión (0,74) y otro para la sensibilidad (0,79)²¹. La validación del MSCEIT en población española encontró un alfa de 0,80³¹. La validación española del RMET encontró un alfa de 0,56³⁰. Aunque el Test de Caras obtuvo un alfa ligeramente inferior a la mayoría de los tests descritos anteriormente es un valor que puede considerarse adecuado para su uso.

La literatura científica que informe de consistencia interna y estabilidad test-retest es muy limitada. La fiabilidad test-retest del JACBART está entre 0,44 para la emoción de ira y el 0,72 para la tristeza (calculada con t de Student), lo que sugiere efecto aprendizaje¹⁷. Hasta donde se sabe, los únicos datos de fiabilidad test-retest en un test de reconocimiento facial de emociones en población española provienen del RMET. Utilizando en método Bland-Altman, se encontró un índice de 0,63²⁹.

En nuestro caso, encontramos una estabilidad test-retest moderada pero significativa

($r = 0,372$, $p = 0,088$). Para evaluarla más exhaustivamente, calculamos los índices Guttman Lambda 3 para cada ítem, ya que es un enfoque más apropiado para ítems binarios. Este índice es equivalente al alfa de Cronbach y puede interpretarse de la misma manera³⁷. En nuestra muestra, 17 de los ítems obtuvieron un valor superior a 0,6, lo que nos permite concluir que el Test de Caras tiene una excelente estabilidad test-retest. La discrepancia entre el valor obtenido usando el test completo y el test-retest ítem a ítem puede ser debida a un menor tamaño de la muestra en el retest.

El Test de Caras fue diseñado usando un modelo de un factor, por lo que fue la única solución que se exploró. Nuestros resultados apenas alcanzaron valores que asuman el ajuste a un único factor. Futuros estudios con este test deberían probar más soluciones y confirmar nuestros resultados.

Puesto que, a pesar de nuestras búsquedas, no tenemos constancia de más validaciones de este test, futuras investigaciones deberían reportar sus propiedades psicométricas y sus valores normativos, lo que nos permitirá alcanzar nuevas conclusiones.

No encontramos diferencias de género. Este es un hallazgo sorprendente, ya que se ha encontrado una pequeña pero consistente superioridad femenina en tests de reconocimiento facial de emociones^{18,31,39}. Lyusin *et al.* (2016) tampoco encontraron diferencias de género en su test, consistente en 7 ítems. En su caso, los autores lo atribuyen a un paradigma más ecológico que podría estar neutralizando la ventaja de las mujeres²¹. Es posible que una posible ventaja femenina solo pueda ser detectada con tests más largos.

Cabe remarcar que un estudio transcultural en el que se comparó el rendimiento de una muestra brasileña y otra francesa en el FERT no encontró diferencias de género en la muestra brasileña, pero sí en la francesa²⁸. Estos hallazgos ponen de manifiesto que algunos aspectos de la interacción entre el reconocimiento facial de emociones y las variables demográficas pueden no ser estables en diferentes culturas. En este caso, los autores del mencionado estudio argumentan que los estereotipos de género y la igualdad de género pueden influenciar el reconocimiento facial de emociones²⁸.

De la misma manera, tampoco encontramos la influencia de la edad que se ha reportado previamente en la literatura^{26,27}. Sin embargo, nuestra muestra está agrupada en torno a la mediana edad y carecemos de suficientes participantes jóvenes y de avanzada edad, por lo que estos resultados eran esperables.

No encontramos diferencias significativas entre distintos niveles académicos y desempeño en el test, lo que no es consistente con literatura previa^{16,28}. De acuerdo con Lind-

quist (2014), la memoria semántica tiene un rol crucial en la denominación de emociones independientemente de su valencia⁴⁰, por lo que especulamos que, a pesar de que un mayor nivel educativo puede ayudar a acceder a un vocabulario emocional más rico, el Test de Caras no representa un reto semántico.

De acuerdo con nuestros resultados, el Test de Caras presenta unas adecuadas propiedades psicométricas, presenta efecto techo y no es sensible a factores demográficos que han sido identificados en la literatura como importantes por su influencia en el reconocimiento facial de emociones. Estas desventajas podrían impedir la habilidad del test de detectar un deterioro sutil en el reconocimiento facial de emociones. A pesar de ello, debido a que la población general normalmente alcanza un efecto techo, es probable que puntuaciones inferiores a una desviación estándar realmente reflejen déficits en reconocimiento facial de emociones. No obstante, recomendamos que esta interpretación sea tomada con prudencia hasta que futuros estudios examinen la sensibilidad-especificidad de este test en diferentes poblaciones clínicas.

Nuestro trabajo debe ser interpretado ante ciertas fortalezas y limitaciones:

En lo que refiere a fortalezas, esta es la primera validación publicada para el Test de Caras y una de las primeras validaciones de un test de reconocimiento facial de emociones para población española. Este trabajo ofrece datos valiosos para su uso clínico y de investigación y como referencia para otros autores que desarrollen datos normativos en medidas de reconocimiento facial de emociones en población española.

Además, conseguimos reclutar personas de todo el territorio español, lo que ofrece una buena validez a nuestros resultados.

En cuanto a limitaciones, debemos recalcar que incluso si una traducción es precisa, las diferencias culturales y la familiaridad con las palabras puede alterar la dificultad del ítem o inducir diferentes respuestas a un test. Esto es de particular importancia, ya que el presente instrumento fue desarrollado para población británica y haberlo adaptado al español puede haber influenciado las respuestas de los participantes. Este efecto puede haber sido más pronunciado en los diez ítems que representan estados mentales complejos.

Además, este test es corto y exhibe la máxima intensidad de la emoción, lo que es probable que haya disminuido la habilidad del test para detectar un desempeño excelente o un deterioro sutil.

A pesar de que usar una encuesta online nos permitió reclutar muestra de diferentes regiones de España, perder el contacto presencial con las personas participantes puede haber influido nuestros resultados.

Finalmente, no pudimos reclutar una muestra equilibrada en género, edad y nivel educativo. Nuestra muestra de mujeres fue lo suficientemente grande como para detectar diferencias, no así la variabilidad en edad y nivel educativo, ya que sobrerrepresenta a personas con mayor nivel educativo.

A pesar de sus limitaciones, consideramos que el presente trabajo ofrece un recurso valioso para personal investigador y clínico en España. Futuros estudios podrían examinar la sensibilidad-especificidad en diferentes poblaciones o en sujetos de mayor edad y menor nivel educativo.

Agradecimientos. Agradecemos a la Cátedra UAM-ASI-SA (Gestión Sanitaria y Economía de la Salud) el financiamiento para este proyecto. Agradecemos a CIBERSAM (Centro de Investigación Biomédica en Red – Salud Mental) por haber contribuido con fondos a este proyecto. Agradecemos su participación a todas las personas voluntarias que ofrecieron su tiempo para participar en este proyecto.

Declaración de intereses. Ninguna persona en la autoría declara tener conflictos de interés.

BIBLIOGRAFÍA

1. Christensen KJ, Riley BE, Heffernan KA, Love SB, Sta. Maria MEM. Facial Recognition Test in the Elderly: Norms, Reliability and Premorbid Estimation. *Clin Neuropsychol (Neuropsychology, Dev Cogn Sect D)* [Internet]. 2002 Feb 1 [cited 2017 Nov 6];16(1):51–6. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/11992226>
2. Kessels RPC, Montagne B, Hendriks AW, Perrett DI, de Haan EHF. Assessment of perception of morphed facial expressions using the Emotion Recognition Task: normative data from healthy participants aged 8–75. *J Neuropsychol* [Internet]. 2014 Mar [cited 2016 Sep 27];8(1):75–93. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23409767>
3. Rodger H, Vizioli L, Ouyang X, Caldara R. Mapping the development of facial expression recognition. *Dev Sci*. 2015;18(6):926–39.
4. Kohler CG, Walker JB, Martin EA, Healey KM, Moberg PJ. Facial emotion perception in schizophrenia: a meta-analytic review. *Schizophr Bull* [Internet]. 2010 Sep [cited 2016 Oct 6];36(5):1009–19. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19329561>
5. Samamé C, Martino DJ, Strejilevich SA. Social cognition in euthymic bipolar disorder: systematic review and me-

- ta-analytic approach. *Acta Psychiatr Scand* [Internet]. 2012 Apr [cited 2016 Oct 13];125(4):266–80. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22211280>
6. Pringle A, Harmer CJ, Cooper MJ. Biases in emotional processing are associated with vulnerability to eating disorders over time [Internet]. *Eating Behaviors* Jan, 2011 p. 56–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21184974>
 7. Dalili MN, Penton-Voak IS, Harmer CJ, Munafò MR. Meta-analysis of emotion recognition deficits in major depressive disorder. *Psychol Med* [Internet]. 2015 Apr [cited 2016 Oct 6];45(6):1135–44. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25395075>
 8. Dalrymple KA, Gomez J, Duchaine B, Anastasi J, Rhodes M, Wright D, et al. The Dartmouth Database of Children's Faces: Acquisition and Validation of a New Face Stimulus Set. Urgesi C, editor. *PLoS One* [Internet]. 2013 Nov 14 [cited 2016 Aug 18];8(11):e79131. Available from: <http://dx.plos.org/10.1371/journal.pone.0079131>
 9. Ekman P, Friesen W V. *Pictures of Facial Affect* [Internet]. Consulting Psychologists Press, editor. Palo Alto, CA; 1976 [cited 2017 Apr 18]. Available from: https://books.google.es/books/about/Pictures_of_Facial_Affect.html?id=gbfMSgAACAAJ&redir_esc=y
 10. Langner O, Dotsch R, Bijlstra G, Wigboldus DHJ, Hawk ST, van Knippenberg A. Presentation and validation of the Radboud Faces Database. *Cogn Emot* [Internet]. 2010 Dec [cited 2016 Nov 10];24(8):1377–88. Available from: <http://www.tandfonline.com/doi/abs/10.1080/02699930903485076>
 11. Marchewka A, Żurawski Ł, Jednoróg K, Grabowska A. The Nencki Affective Picture System (NAPS): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database. *Behav Res Methods* [Internet]. 2014 Jun [cited 2017 Nov 6];46(2):596–610. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23996831>
 12. Vaiman M, Wagner MA, Caicedo E, Pereno GL. Development and validation of an Argentine set of facial expressions of emotion. *Cogn Emot* [Internet]. 2017 Feb 17 [cited 2017 Nov 6];31(2):249–60. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26479048>
 13. Olszanowski M, Pochwatko G, Kuklinski K, Scibor-Rylski M, Lewinski P, Ohme RK. Warsaw set of emotional facial expression pictures: a validation study of facial display photographs. *Front Psychol* [Internet]. 2015 Jan 5 [cited 2017 Nov 6];5:1516. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25601846>
 14. Tottenham N, Tanaka JW, Leon AC, McCarry T, Nurse M, Hare TA, et al. The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Res* [Internet]. 2009 Aug 15 [cited 2019 May 2];168(3):242–9. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0165178108001480>
 15. Pinkham AE, Penn DL, Green MF, Buck B, Healey K, Harvey PD. The Social Cognition Psychometric Evaluation Study: Results of the Expert Survey and RAND Panel. *Schizophr Bull* [Internet]. 2014 Jul [cited 2016 Aug 1];40(4):813–23. Available from: <https://academic.oup.com/schizophreniabulletin/article-lookup/doi/10.1093/schbul/sbt081>
 16. Dodich A, Cerami C, Canessa N, Crespi C, Marcone A, Arpone M, et al. Emotion recognition from facial expressions: a normative study of the Ekman 60-Faces Test in the Italian population. *Neurol Sci* [Internet]. 2014 Jul 19 [cited 2018 Feb 22];35(7):1015–21. Available from: <http://link.springer.com/10.1007/s10072-014-1631-x>
 17. Matsumoto D, LeRoux J, Wilson C, Raroque J, Kooen K, Ekman P, et al. A New Test to Measure Emotion Recognition Ability: Matsumoto and Ekman's Japanese and Caucasian Brief Affect Recognition Test (JACBART). *J Nonverbal Behav* [Internet]. 2000 [cited 2019 May 20];24(3):179–209. Available from: <http://link.springer.com/10.1023/A:1006668120583>
 18. Passarelli M, Masini M, Bracco F, Petrosino M, Chiorri C. Development and validation of the Facial Expression Recognition Test (FERT). *Psychol Assess* [Internet]. 2018 Nov [cited 2019 Apr 16];30(11):1479–90. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/30024180>
 19. Bryson G, Bell M, Lysaker P. Affect recognition in schizophrenia: a function of global impairment or a specific cognitive deficit. *Psychiatry Res* [Internet]. 1997 Jul 4 [cited 2017 Apr 24];71(2):105–13. Available from: <http://www.sciencedirect.com.sire.ub.edu/science/article/pii/S0165178197000504?via%3Dihub>
 20. Baron-Cohen S, Wheelwright S, Jolliffe and T. Is There a "Language of the Eyes"? Evidence from Normal Adults, and Adults with Autism or Asperger Syndrome. *Vis cogn* [Internet]. 1997 Sep [cited 2016 Oct 26];4(3):311–31. Available from: <http://www.tandfonline.com/doi/abs/10.1080/713756761>
 21. Lyusin, Dmitry; Ovsyannikova V. Measuring two aspects of emotion recognition ability: Accuracy vs. sensitivity. *Learn Individ Differ* [Internet]. 2016 Dec 1 [cited 2019 May 20];52:129–36. Available from: <https://www.sciencedirect-com.sire.ub.edu/science/article/pii/S1041608015001090>
 22. McDonald S, Flanagan S, Rollins J, Kinch J. TASIT: A new clinical tool for assessing social perception after traumatic brain injury. *J Head Trauma Rehabil* [Internet]. [cited 2018 Nov 20];18(3):219–38. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/12802165>
 23. McDonald S, Honan C, Allen SK, El-Helou R, Kelly M, Kumfor F, et al. Normal adult and adolescent performance on TASIT-S, a short version of The Assessment of Social Inference Test. *Clin Neuropsychol* [Internet]. 2018 May 19 [cited 2019 Sep 3];32(4):700–19. Available from: <https://www.tandfonline.com/doi/full/10.1080/13854046.2017.1400106>

24. McClure EB. A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents. *Psychol Bull* [Internet]. 2000 May [cited 2016 Sep 27];126(3):424–53. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/10825784>
25. Thompson AE, Voyer D. Sex differences in the ability to recognise non-verbal displays of emotion: A meta-analysis. *Cogn Emot* [Internet]. 2014 Oct 3 [cited 2016 Aug 8];28(7):1164–95. Available from: [http://www.tandfonline.com/doi/abs/10,1080/02699931.2013,875889](http://www.tandfonline.com/doi/abs/10.1080/02699931.2013.875889)
26. Lawrence K, Campbell R, Skuse D. Age, gender, and puberty influence the development of facial emotion recognition. *Front Psychol* [Internet]. 2015 Jun 16 [cited 2016 Aug 16];6:761. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26136697>
27. Ruffman T, Henry JD, Livingstone V, Phillips LH. A meta-analytic review of emotion recognition and aging: Implications for neuropsychological models of aging. *Neurosci Biobehav Rev*. 2008;32(4):863–81.
28. de Souza LC, Bertoux M, de Faria ÂRV, Corgosinho LTS, Prado AC de A, Barbosa IG, et al. The effects of gender, age, schooling, and cultural background on the identification of facial emotions: a transcultural study. *Int Psychogeriatrics* [Internet]. 2018 Dec 25 [cited 2019 Mar 13];30(12):1861–70. Available from: https://www.cambridge.org/core/product/identifier/S1041610218000443/type/journal_article
29. Fernández-Abascal EG, Cabello R, Fernández-Berrocal P, Baron-Cohen S. Test-retest reliability of the "Reading the Mind in the Eyes" test: a one-year follow-up study. *Mol Autism* [Internet]. 2013 Sep 11 [cited 2018 Apr 4];4(1):33. Available from: <http://molecularautism.biomedcentral.com/articles/10,1186/2040-2392-4-33>
30. Redondo I, Herrero-Fernández D. Validation of the Reading the Mind in the Eyes Test in a healthy Spanish sample and women with anorexia nervosa. *Cogn Neuropsychiatry* [Internet]. 2018 Jul 4 [cited 2019 May 20];23(4):201–17. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29635964>
31. Extremera N, Fernández-Berrocal P, Salovey P. Spanish version of the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT). Version 2.0: reliabilities, age and gender differences. *Psicothema* [Internet]. 2006 [cited 2018 Sep 25];18 Suppl:42–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17295956>
32. Flores I, Casaletto KB, Marquine MJ, Umlauf A, Moore DJ, Mungas D, et al. Performance of Hispanics and Non-Hispanic Whites on the NIH Toolbox Cognition Battery: the roles of ethnicity and language backgrounds. *Clin Neuropsychol* [Internet]. 2017 May 19 [cited 2019 Sep 3];31(4):783–97. Available from: <https://www.tandfonline.com/doi/full/10,1080/13854046.2016,1276216>
33. Muñiz J, Elosua P, Hambleton RK, International Test Commission. [International Test Commission Guidelines for test translation and adaptation: second edition]. *Psicothema* [Internet]. 2013 [cited 2018 Mar 2];25(2):151–7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23628527>
34. Baron-Cohen S, Jolliffe T, Mortimore C, Robertson M. Another advanced test of theory of mind: evidence from very high functioning adults with autism or asperger syndrome. *J Child Psychol Psychiatry* [Internet]. 1997 Oct [cited 2016 Oct 25];38(7):813–22. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9363580>
35. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing [Internet]. Vienna, Austria: ISBN 3–900051–07–0; 2012. Available from: <https://www.r-project.org/>
36. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA* [Internet]. 2013 Nov 27 [cited 2017 Jul 6];310(20):2191–4. Available from: <http://jama.jamanetwork.com/article.aspx?doi=10,1001/jama.2013.281053>
37. Guttman L. A basis for analyzing test-retest reliability. *Psychometrika* [Internet]. 1945 Dec [cited 2017 Mar 30];10(4):255–82. Available from: <http://link.springer.com/10,1007/BF02288892>
38. Kessels RPC, Montagne B, Hendriks AW, Perrett DI, de Haan EHF. Assessment of perception of morphed facial expressions using the Emotion Recognition Task: Normative data from healthy participants aged 8–75. *J Neuropsychol* [Internet]. 2014 Mar [cited 2017 Nov 6];8(1):75–93. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23409767>
39. Bora E, Pantelis C, Addington J, Addington D, Baron-Cohen S, Wheelwright S, et al. Social cognition in schizophrenia in comparison to bipolar disorder: A meta-analysis. *Schizophr Res* [Internet]. 2016 Aug [cited 2016 Jul 25];175(1–3):72–8. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0920996416301670>
40. Lindquist KA, Gendron M, Barrett LF, Dickerson BC. Emotion perception, but not affect perception, is impaired with semantic memory loss. *Emotion* [Internet]. 2014 Apr [cited 2019 May 21];14(2):375–87. Available from: <http://doi.apa.org/getdoi.cfm?doi=10,1037/a0035293>