

## Artículo original

# Estructura Interna y Validez del Reading the Mind in the Eyes Test Versión Niños en una Muestra Colombiana de Estudiantes de Secundaria<sup>☆</sup>



Cristian Ávila-Rincón<sup>a,\*</sup>, Jenny Benitez-Murillo<sup>a</sup> y Steven Rodríguez-Portillo<sup>b,c</sup>

<sup>a</sup> Psychology program, Health Sciences Faculty, Pedagogical and Technological University of Colombia, Tunja, Colombia

<sup>b</sup> Health and Clinical Psychology Research Group (GIPCS), Health Sciences Faculty, Pedagogical and Technological University of Colombia, Tunja, Colombia

<sup>c</sup> Basic Psychological Processes Laboratory (LAPROPSIC), Health Sciences Faculty, Pedagogical and Technological University of Colombia, Tunja, Colombia

### INFORMACIÓN DEL ARTÍCULO

#### Historia del artículo:

Recibido el 20 de julio de 2021

Aceptado el 4 de julio de 2022

On-line el 26 de agosto de 2022

#### Palabras clave:

Cognición social

Teoría de la mente

Validez

Empatía

### RESUMEN

**Introducción y objetivos:** La cognición social es un conjunto de procesos relacionados con la interacción interpersonal, la adaptación social y toda una serie de comportamientos sociales complejos. Las alteraciones de estos procesos son una característica clave de muchos trastornos psiquiátricos, por lo que la adaptación y la validación de las herramientas de evaluación del desempeño de la cognición social son fundamentales para un diagnóstico y un tratamiento adecuados. Un instrumento de evaluación muy popular es el Reading the Mind in the Eyes Test (RMET), que se utiliza para evaluar la capacidad individual en cuanto a habilidades de cognición social. A pesar de su notable popularidad, hasta la fecha no existen registros sobre sus características psicométricas en población infantil de habla hispana. Por ello, el objetivo de esta investigación es evaluar la consistencia interna, la validez convergente y la estructura factorial del RMET en población infantil colombiana.

**Métodos:** Nuestra muestra consistió en estudiantes escolares (n=509; 315 varones y 194 mujeres) de sexto a noveno grado con edades entre los 10 y los 16 años.

**Resultados:** El RMET tiene una consistencia interna aceptable, una validez convergente moderada y un aceptable ajuste a una estructura factorial unidimensional que podría mejorar utilizando una versión abreviada. Además, no se encontraron diferencias significativas entre los grupos de sexo y edad, y las puntuaciones medias de rendimiento de cada prueba parecen ser similares a las de las personas de entornos culturales afines (España o Argentina).

<sup>☆</sup> Con base en la tesis: «Propiedades psicométricas de dos instrumentos que evalúan cognición social en niños y adolescentes» (2020). Universidad Pedagógica y Tecnológica de Colombia.

\* Autor para correspondencia.

Correo electrónico: [Cristian.avila01@uptc.edu.co](mailto:Cristian.avila01@uptc.edu.co) (C. Ávila-Rincón).

<https://doi.org/10.1016/j.rcp.2022.07.002>

0034-7450/© 2022 Asociación Colombiana de Psiquiatría. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

*Conclusiones:* El RMET posee una fiabilidad y una validez aceptables y, por lo tanto, es una prueba adecuada para diferenciar las habilidades de teoría de la mente en población con desarrollo típico. Se debe investigar sus propiedades psicométricas en muestras clínicas y otros grupos etarios.

© 2022 Asociación Colombiana de Psiquiatría. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

## Internal Structure and Validity of the Reading the Mind in the Eyes Test Child V ERSION in a Sample of Colombian High Schoolers

### A B S T R A C T

#### Keywords:

Social cognition  
Theory of Mind  
Validity  
Empathy

*Introduction and objectives:* Social cognition is a broad set of processes related to interpersonal interaction, social adjustment, and a whole array of complex social behaviours. Disturbances in these processes are a key feature of many psychiatric disorders, therefore the adaptation and validation of assessment tools to measure social cognition performance are critical for correct diagnosis and treatment. One popular assessment tool is the Reading the Mind in the Eyes Test (RMET), which is used to measure individual ability regarding social cognition skills. Despite its remarkable popularity, to date there is no record about its psychometric features in Spanish-speaking children. Therefore, the aim of this research was to assess the internal consistency, convergent validity, and factorial structure of the RMET in children.

*Methods:* Our sample consisted of school students ( $n = 509$ ; 315 males, and 194 females) from sixth to ninth grade with ages ranging from 10 to 16 years old.

*Results:* The RMET has acceptable internal consistency, moderate convergent validity, and acceptable adjustment to a unidimensional factorial structure that could be improved by using a shortened version. No significant differences between sex and age groups were found, average performance scores of each test seem to be similar to those from akin cultural backgrounds (Spain or Argentina).

*Conclusions:* The RMET has acceptable reliability and validity and is therefore a suitable test for differentiating theory of mind skills in typically developing populations. Its psychometric properties should be investigated in clinical samples and other age groups.

© 2022 Asociación Colombiana de Psiquiatría. Published by Elsevier España, S.L.U. All rights reserved.

## Introducción

La cognición social (CS) hace referencia a la habilidad de procesar información social relevante con la finalidad de generar y garantizar un adecuado comportamiento social<sup>1–3</sup>. Entre los enfoques de estudio de la CS, algunos proponen que esta trabaja a partir de 2 sistemas complementarios; el primero de ellos es el sistema implícito, también conocido como cognición social de bajo nivel, el cual funciona de forma automática, espontánea, inconsciente y veloz (p. ej., atribuir estados mentales en otros); el otro es el sistema explícito, llamado también cognición social de alto nivel, que se caracteriza por ser consciente, lento y reflexivo<sup>4</sup>. Además de cómo trabaja la CS, también es relevante mencionar que está compuesta por una variedad de procesos, de entre los que se puede destacar: la teoría de la mente (ToM), la percepción social y el procesamiento emocional<sup>5,6</sup>.

La ToM hace referencia a la capacidad para atribuir estados mentales (tales como deseos, creencias o emociones, entre otros) de terceros y usar tal información para entender qué

sabe/piensa la otra persona y de esta manera predecir su comportamiento. Según Shamay-Tsoory et al.<sup>7</sup>, la ToM está conformada por 2 componentes, a saber: ToM emocional o afectiva y ToM cognitiva. A grandes rasgos, la emocional es la capacidad de inferir lo que la persona siente, mientras que el componente cognitivo se enfoca más en la inferencia de creencias, motivaciones y pensamientos<sup>8</sup>. Como mencionan Brown et al.<sup>9</sup>, los conceptos de empatía y ToM emocional suelen confundirse; sin embargo, a diferencia de la ToM, la empatía hace referencia a la habilidad de las personas de entender y reaccionar a las emociones de otros.

A lo largo de los años, se ha desarrollado gran cantidad de instrumentos para evaluar la ToM (para una revisión, véase Beaudoin et al.<sup>10</sup>); no obstante, una de las más destacadas y empleada con mayor frecuencia es la prueba *Reading the Mind in the Eyes Test* (RMET), también conocida como test de las miradas. El RMET evalúa principalmente la capacidad de reconocer estados mentales en otros y la habilidad para entender emociones simples.

La adquisición de la ToM en los niños tiene un papel muy importante, puesto que los déficits en su desarrollo están aso-

**Tabla 1 – Investigaciones previas con el RMET versión niños**

Investigación	Muestra	Edades (años)	Media ± DE	Confiabilidad
Müller et al., 2013	N=596 (V, 303; M, 293)	13-15	17,8 ± 3,5	$\alpha = 0,550$
Van der Meulen et al., 2017	N=718 (V, 343; M, 374)	6-9	16,98	ICC = 0,469
Rueda et al., 2013	N=354 (V, 163; M, 191)	14-16	18,51 ± 3,93	NA
Vogindroukas et al., 2014	N=80 (V, 52; M, 28)	9-10	16,4 ± 3,01	$\alpha = 0,687$
Girli, 2014	N=235 (V, 136; M, 99)	6-16		$\alpha = 0,720$

DE: desviación estándar; RMET: *Reading the Mind in the Eyes*; M: mujeres; V: varones.

ciados generalmente con una serie de trastornos (p. ej., del espectro de la esquizofrenia, del espectro autista o problemas del lenguaje) y con una incorrecta interpretación de las diferentes claves sociales, lo que puede llevar también a un comportamiento social disfuncional.

En Colombia, la investigación en cuanto a ToM en niños ha sido casi nula, y en el caso de evaluación con el test de las miradas no existen reportes previos de su aplicación. Una de las probables razones de que dichas investigaciones no se hayan realizado es la falta de validación del instrumento en población infantil colombiana. Por ello el presente estudio tiene como objetivo evaluar la consistencia interna, la validez convergente y la estructura factorial del RMET en población infantil colombiana.

A pesar de ser un instrumento muy usado en ámbitos como el clínico y el educativo, el RMET en su versión para niños tiene poca evidencia sobre sus características psicométricas en el mundo. Tras una revisión bibliográfica, se encontraron en total 5 estudios donde se describen las propiedades psicométricas<sup>11-15</sup>.

En términos de confiabilidad, los datos se obtuvieron principalmente a través del alfa de Cronbach, con valores entre 0,55 y 0,72 (tabla 1).

En cuanto a la validez del instrumento, existe variedad en los resultados, dependiendo del tipo de análisis de validez que se utilizara; por ejemplo, en el caso de la validez de constructo, la mayoría de las investigaciones usan el análisis factorial<sup>11,16-18</sup>.

Sobre la validez convergente, Vellante et al.<sup>18</sup> muestran en su revisión que, de un total de 24 investigaciones que trabajaron con propiedades psicométricas del test de las miradas (principalmente estudios con muestra de adultos), 6 muestran correlación con el cuestionario de empatía (EQ) con valores de  $r$  entre 0,27 y 0,56, así como correlación con la *Emotion Recognition Task* (ERT) ( $r = 0,38$ ). Por último, una de las investigaciones incluidas en la revisión reporta que no hubo correlación entre el RMET y el índice de reactividad interpersonal (IRI).

Asimismo Olderbak et al.<sup>17</sup> reportan evidencia de validez convergente entre el test de las miradas y otros instrumentos que evalúan ToM (*Faux pas test*), al igual que entre dicho instrumento y pruebas que miden empatía. No obstante, las correlaciones con este último constructo (empatía) son variadas, en ocasiones negativas y en ocasiones positivas, pero siempre correlaciones débiles.

Teniendo en cuenta lo mencionado, el objetivo de la presente investigación es evaluar la consistencia interna, la validez convergente y la estructura factorial del RMET en población infantil colombiana.

**Tabla 2 – Características sociodemográficas de la muestra**

	(n)	%
<b>Sexo</b>		
Mujeres	194	38,12
Varones	315	61,88
<b>Zona de residencia</b>		
Urbana	468	91,94
Rural	41	8,06
<b>Estrato socioeconómico</b>		
Estrato 1	91	17,87%
Estrato 2	372	73,08%
Estrato 3	46	9,05%
<b>Estado civil de los padres</b>		
Soltero	18	3,54
Casados	253	49,70
Unión libre	107	21,02
Separado	113	22,22
Divorciado	9	1,76
Viudo	9	1,76
<b>Grado</b>		
Sexto	69	33,01
Séptimo	127	24,95
Octavo	101	19,84
Noveno	212	41,65

## Métodos

### Participantes

Se seleccionó una muestra no probabilística de 509 niñas y niños, estudiantes de un colegio público del departamento de Boyacá (Colombia). Las edades estaban entre 10 y 16 años, con una media de  $13,10 \pm 1,30$  años, y cursaban entre sexto y noveno grado (tabla 2). Con el fin de contar con la participación de los menores, se pidió el asentimiento de los niños y la autorización por medio del consentimiento informado diligenciado por el acudiente.

Se realizó una entrevista con los padres de los participantes, el docente de apoyo de la sede y los docentes directores de curso de cada grado, con el propósito de detectar cualquier condición de salud previa o actual. Se establecieron como criterios de exclusión: cualquier condición de salud que pudiera afectar a la participación de los menores o un puntaje por debajo de la media en una prueba de cribado de desempeño académico. Los criterios de inclusión fueron: desear participar voluntariamente en el estudio y poseer la autorización firmada por el acudiente del menor.



1. Amigable 2. Triste 3. Sorprendido 4. Aburrido



1. Celso 2. Asutado 3. Amigable 4. Sorprendido

**Figura 1 – Reading the Mind in the Eyes (RMET), versión niños.**

El desarrollo del proceso de investigación se enmarcó dentro de los parámetros establecidos por el Colegio Colombiano de Psicólogos (COLPSIC), el Código de Ética y Bioética del Psicólogo en Colombia y el Código de Ética de la Asociación Médica Mundial (Declaración de Helsinki).

### Instrumentos

#### Reading the Mind in the Eyes

El TMET versión infantil (RMET-C)<sup>19</sup> es una prueba que evalúa el reconocimiento de estados mentales complejos a través la región ocular del rostro humano. Consta de 28 fotografías con diferentes expresiones y 4 opciones de respuesta, de las cuales solo una reproduce el estado emocional o mental correcto representado por la imagen; cada elemento se puntúa como correcto o incorrecto.

Para este estudio se utilizó la versión en español del test de las miradas, que cuenta con 28 fotografías y fue traducida por Pineda-Alhucema et al.<sup>20</sup>. En esta prueba, cada participante debe elegir entre cuatro posibles opciones de respuesta para cada fotografía la palabra que mejor describa lo que la persona de la imagen está pensando o sintiendo (figura 1). Se otorga 1 punto por cada respuesta correcta, hasta un máximo de 28 puntos.

Aunque los autores de la prueba no indicaron un punto de corte para identificar a individuos con una posible alteración, reportes de estudios anteriores<sup>21</sup> indican que un desempeño por debajo de 2 desviaciones estándar de la media del grupo de referencia indicarían alteración en la capacidad de reconocer emociones y estados mentales complejos.

#### Cambridge Mindreading Face-Voice Battery

La batería Cambridge Mindreading Face-Voice para niños (CAM-C) mide el reconocimiento de emociones y estados mentales complejos a través de la expresión facial y vocal. La tarea facial incluye 27 clips de vídeo cortos de actores que representan una emoción (figura 2). La tarea vocal consiste en 27 audios con verbalizaciones cortas pronunciadas en una entonación emo-



1. Triste 2. Enojado 3. Amigable 4. Aburrida



1. Molesto 2. Amable 3. Cansado 4. Sorprendido

**Figura 2 – CAM, versión niños.**

cional particular. Para cada clip de audio y vídeo, se muestran 4 palabras con emociones numeradas del 1 al 4, pero solo una reproduce el estado emocional o mental correcto que el audio o vídeo representa<sup>22</sup>.

Según los autores del instrumento<sup>22</sup>, en niños se evidenciaría déficit en el reconocimiento de emociones cuando los resultados están 1,15 desviaciones estándar por debajo de la media del grupo de control.

#### Interpersonal Reactivity Index

El IRI es una medida autoaplicada que evalúa varios componentes de la empatía desde una perspectiva multidimensional. Está conformada por 28 ítems que siguen un formato de respuesta de tipo Likert de 5 puntos y van desde: «No me describe bien» hasta «Me describe muy bien». Posee 4 subescalas, cada una de ellas compuesta por 7 ítems diferentes; cada subescala describe un componente específico de la habilidad empática, a saber: toma de perspectiva, fantasía, preocupación empática y angustia personal.

Esta medida ofrece un indicador cuantitativo de las diferencias individuales en 2 constructos generales: la empatía cognitiva (conformada por las subescalas de toma de perspectiva y fantasía), que es la capacidad de comprender la perspectiva de la otra persona, y la empatía afectiva (compuesta por preocupación empática y angustia personal), que es la habilidad de reaccionar apropiadamente a los estados emocionales de otros<sup>23</sup>. Cuanto mayor sea el puntaje obtenido, mayor nivel del rasgo posee un individuo. Se utilizó la versión infantil española, que presenta características psicométricas similares a la versión original.

#### Procedimiento

Se programaron sesiones de 45 minutos para que los participantes tomaran las pruebas computarizadas, un investigador entrenado en el procedimiento proporcionaba las instrucciones a grupos de 20 menores aproximadamente. Durante todo

el proceso esta misma persona estuvo presente para responder a cualquier duda o facilitar la hoja de definiciones en caso de que los menores no estuvieran familiarizados con alguna palabra presentada en la prueba. Se proporcionaron auriculares para tomar la tarea vocal de CAM-C. El orden de presentación de las tareas fue RMET, CAM-C e IRI y todas las pruebas utilizadas en este estudio se presentaron utilizando el software OpenSesame<sup>24</sup>.

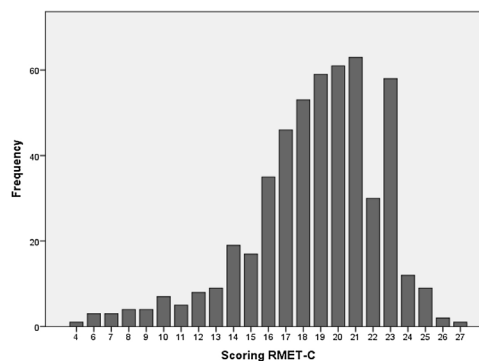
### Análisis de datos

El análisis de los datos siguió un proceso de 3 pasos. En primer lugar, se calcularon los estadísticos descriptivos de todas las pruebas utilizadas; luego, para determinar las relaciones entre las pruebas administradas, se calcularon las matrices de correlación entre las puntuaciones de las respuestas correctas de todas las pruebas. Teniendo en cuenta las investigaciones anteriores, se esperó una correlación débil o nula entre la medida de empatía (IRI) y la medida de la ToM (RMET), pero una moderada entre ambas pruebas de ToM (RMET-CAM). Se calculó y mostró la distribución de las respuestas seleccionadas, también se exploraron las diferencias en las puntuaciones con respecto a la edad, el sexo y el grado de los estudiantes, por lo que se utilizaron ANOVAS de 1 y 2 factores para establecer estas posibles diferencias.

En segundo lugar, en todos los ítems se determinaron los índices de dificultad (Pi) y de discriminación (Di) siguiendo el modelo lineal de la teoría clásica de los tests para caracterizar su rendimiento<sup>25</sup>. Los coeficientes alfa de Cronbach y omega de McDonald's se utilizaron a fin de determinar los índices de consistencia interna y saturación factorial<sup>26-28</sup>.

En tercer lugar, para determinar la estructura factorial de los datos se realizó un análisis factorial confirmatorio (AFC) y se estableció una solución factorial unidimensional para el test RMET. No obstante, algunos estudios en adultos han reportado una solución de 3 factores basada en la valencia de los ítems<sup>29</sup>. Hasta donde se sabe, ningún estudio en niños ha especificado antes dicho modelo. El AFC se basó en una matriz de correlación tetracórica; se utilizó el método de análisis paralelo óptimo para determinar la estructura factorial<sup>30</sup> y se eligió un método de mínimos cuadrados ponderados diagonales (DWLS) como método de extracción debido a las propiedades de la matriz utilizada y la distribución de los datos<sup>31</sup>. Además, se utilizaron los siguientes índices de ajuste para evaluar la calidad del modelo: el error cuadrático medio de aproximación (RMSEA), el índice de ajuste no normalizado (NNFI/TLI), el índice de ajuste comparativo (CFI) y el índice de bondad de ajuste (GFI). También se incluyó la razón  $\chi^2/gf$  y los estadísticos de resumen para los residuos ajustados mediante el residuo cuadrático medio ponderado (WRMR). Se consideraron los puntos de corte estándar cuando se estimó el ajuste del modelo hallado<sup>32</sup>.

Se utilizó el software estadístico JASP<sup>33</sup> para realizar el análisis estadístico descriptivo. El AFC se realizó con el software FACTOR<sup>34</sup> y los coeficientes Di y Pi se calcularon con el paquete sjPlot<sup>35</sup> del software estadístico R<sup>36</sup>.



**Figura 3 – Distribución de las puntuaciones del Reading the Mind in the Eyes Test para niños.**

## Resultados

El intervalo de las puntuaciones del presente estudio fue bastante amplio (4 a 27); la media de la muestra fue de  $18,69 \pm 3,76$  (intervalo de confianza del 95% [IC95%], 18,37-19,02). Como indican los datos y la figura 3, la distribución de las puntuaciones tuvo un ligero sesgo negativo, pero con valores dentro de la franja esperada<sup>37</sup>. Se utilizaron pruebas no paramétricas para el análisis estadístico debido a que, según los análisis de normalidad (prueba de Kolmogorov-Smirnov,  $p < 0,01$ ), las puntuaciones de todas las variables no siguieron una distribución normal.

### Distribución de las respuestas correctas

La elección de respuestas correctas osciló entre el 18,6 y el 87,3%, con una media del  $60,82\% \pm 17,27\%$ ; de los 28 ítems, 18 obtuvieron una elección correcta por al menos el 60% de los participantes. Al analizar la tasa de elección correcta por el sexo de los participantes, la única diferencia relevante se encontró en el ítem 26 (correspondiente a la emoción «nerviosa») favoreciendo a los niños sobre las niñas (niñas, el 38,1% frente al 46,7% de los niños). En los demás ítems las diferencias no fueron superiores al 4%. Solo en 4 ítems las palabras incorrectas tuvieron un mayor porcentaje de elección que la opción correcta; al menos el 30% de los participantes eligieron las palabras incorrectas para estos ítems, que en su mayoría mostraban emociones negativas o neutras. La única diferencia relevante entre niños y niñas en cuanto al porcentaje de elección de la palabra incorrecta se encontró en el ítem 25 (preocupado), más en las niñas que en los niños (el 34,5 frente al 28,9%).

### Análisis de ítems individuales

Los índices de dificultad (Pi) y discriminación (Di) se calcularon mediante los procedimientos habituales<sup>38</sup> con el paquete sjPlot. El Pi medio de los ítems del RMET fue de 0,668 y la media de Di de la misma prueba fue de 0,198. De acuerdo con las clasificaciones estándar, 14 de los 28 ítems se categorizaron como ítems «moderadamente fáciles» ( $0,70 < Pi < 0,90$ ), 9 como ítems

**Tabla 3 – Estadísticos descriptivos y desempeño psicométrico de los ítems del RMET**

Ítem	Individuos que acertaron, %	Media (DE)	IC95%	Alfa de Cronbach	Omega de McDonald	Cargas factoriales	Coefficiente de dificultad (Pi)	Coefficiente de discriminación (Di)
1.	33,1	0,350 ± 0,477	0,30-0,40	0,637	0,662	0,171	0,35	0,039
2.	34,1	0,354 ± 0,479	0,30-0,41	0,645	0,669	-0,084	0,35	-0,041
3.	42,5	0,485 ± 0,500	0,43-0,54	0,629	0,655	0,242	0,49	0,124
4.	76,5	0,790 ± 0,408	0,74-0,84	0,617	0,645	0,346	0,79	0,235
5.	57,9	0,603 ± 0,490	0,55-0,66	0,612	0,642	0,434	0,60	0,271
6.	63,0	0,652 ± 0,477	0,60-0,71	0,618	0,647	0,351	0,65	0,22
7.	72,6	0,768 ± 0,422	0,72-0,82	0,617	0,645	0,381	0,77	0,238
8.	79,8	0,811 ± 0,392	0,77-0,86	0,612	0,639	0,423	0,81	0,304
9.	61,6	0,644 ± 0,479	0,59-0,70	0,617	0,645	0,377	0,64	0,234
10.	18,6	0,609 ± 0,488	0,55-0,66	0,630	0,657	0,113	0,61	0,111
11.	69,3	0,715 ± 0,452	0,66-0,77	0,617	0,645	0,376	0,72	0,231
12.	74,6	0,780 ± 0,415	0,73-0,83	0,609	0,637	0,466	0,78	0,319
13.	65,9	0,760 ± 0,427	0,71-0,81	0,613	0,642	0,379	0,76	0,278
14.	77,1	0,770 ± 0,421	0,72-0,82	0,610	0,638	0,445	0,77	0,313
15.	55,2	0,560 ± 0,497	0,50-0,62	0,637	0,662	0,023	0,56	0,047
16.	75,5	0,802 ± 0,399	0,76-0,85	0,611	0,639	0,438	0,8	0,302
17.	64,2	0,664 ± 0,427	0,61-0,72	0,616	0,646	0,336	0,66	0,238
18.	74,8	0,780 ± 0,415	0,73-0,83	0,616	0,644	0,377	0,78	0,245
19.	59,3	0,615 ± 0,487	0,56-0,67	0,622	0,649	0,299	0,61	0,188
20.	87,3	0,890 ± 0,313	0,85-0,93	0,619	0,645	0,352	0,89	0,234
21.	73,4	0,756 ± 0,430	0,71-0,81	0,618	0,646	0,350	0,76	0,228
22.	73,6	0,760 ± 0,427	0,71-0,81	0,609	0,639	0,404	0,76	0,319
23.	81,4	0,837 ± 0,370	0,79-0,88	0,613	0,639	0,478	0,84	0,298
24.	82,0	0,849 ± 0,359	0,81-0,89	0,617	0,644	0,375	0,85	0,25
25.	36,8	0,369 ± 0,483	0,31-0,42	0,632	0,659	0,061	0,37	0,089
26.	43,6	0,450 ± 0,498	0,39-0,51	0,647	0,670	-0,108	0,45	-0,047
27.	63,8	0,660 ± 0,474	0,61-0,71	0,624	0,653	0,229	0,66	0,165
28.	53,6	0,611 ± 0,488	0,56-0,67	0,630	0,656	0,216	0,61	0,109

«medianamente difíciles» ( $0,51 < \text{Pi} < 0,70$ ) y 5 como ítems «moderadamente difíciles» ( $0,31 < \text{Pi} < 0,50$ ). En cuanto a la distribución de los coeficientes Di, se encontró que al menos 10 ítems del RMET tienen un rendimiento pobre, es decir, un  $\text{Pi} < 0,290$ , mientras que el resto de los ítems tienen un buen rendimiento ( $\text{Di} > 0,30$ ). Estos resultados muestran un desempeño aceptable para casi todos los ítems del RMET (tabla 3), pero indican que algunos deben ser revisados para que tengan un rendimiento adecuado en la muestra presente.

### Consistencia interna y saturación factorial

El total de respuestas correctas se utilizó para determinar las posibles relaciones entre ToM y empatía. En general, la correlación del RMET con las puntuaciones del CAM-C fue positiva, pero moderada ( $r = 0,385$ ;  $p < 0,001$ ), mientras que la relación entre el RMET y el IRI fue nula y estadísticamente no relevante ( $r = -0,041$ ;  $p = 0,385$ ). Se obtuvieron coeficientes de saturación factorial ( $\omega$ ) y de consistencia interna ( $\alpha$ ) para todas las pruebas. Para el RMET la consistencia interna ( $\alpha = 0,630$ ) y la saturación factorial ( $\omega = 0,657$ ) fueron regulares; para el IRI estas medidas fueron buenas ( $\alpha = 0,835$ ;  $\omega = 0,838$ ) y para la prueba CAM-C ( $\alpha = 0,734$ ;  $\omega = 0,739$ ) las medidas fueron aceptables. Al dividir la muestra según el sexo de los participantes, los coeficientes mencionados se mantuvieron similares (niñas,  $\alpha = 0,600$  y  $\omega = 0,638$ ; niños,  $\alpha = 0,648$  y  $\omega = 0,671$ ). La correlación media entre ítems fue de 0,062 y la correlación ítem-prueba osciló entre -0,047 y 0,319.

### Desempeño según variables sociodemográficas

No se encontraron diferencias por sexo al comparar las puntuaciones medias del RMET (niños,  $18,57 \pm 3,87$ ; niñas,  $18,87 \pm 3,57$ ;  $\omega = 0,941$ ;  $p = 0,365$ ;  $r_{\text{tb}} = -0,047$ ). Al analizar otras variables de interés, no se encontraron diferencias estadísticamente relevantes en la puntuación del RMET entre los grupos de edad ( $F(6,502) = 5,276$ ;  $p = 0,624$ ;  $\eta^2 = 0,004$ ). La prueba *post hoc* confirmó este hallazgo (con valores de  $p$  que oscilaron entre 0,060 y 0,485). El mismo resultado se obtuvo para la variable grado, donde no se encontraron diferencias estadísticamente relevantes entre los grupos ( $F(3,505) = 12,630$ ;  $p = 0,830$ ;  $\eta^2 = 0,005$ ). Las pruebas *post hoc* no señalaron ninguna diferencia específica entre los grupos (con valores de  $p$  que variaron entre 0,062 y 0,408). Por último, no se hallaron interacciones entre el sexo y la edad ( $F(9,491) = 6,284$ ;  $p = 0,784$ ;  $\eta^2 = 0,003$ ) ni entre el sexo y el grado ( $F(7,494) = 17,179$ ;  $p = 0,423$ ;  $\eta^2 = 0,005$ ).

### Análisis factorial confirmatorio (AFC)

Se probó una solución inicial de un solo factor mediante el AFC ( $\chi^2 = 627,32$ ;  $\text{gl} = 350$ ;  $p < 0,001$ ;  $\chi^2/\text{gl} = 1,79$ ;  $\text{RMSEA} = 0,037$ ;  $\text{CFI} = 0,836$ ;  $\text{NNFI} = 0,823$ ;  $\text{GFI} = 0,862$ ;  $\text{WRMR} = 0,0572$ ). Los índices de ajuste del modelo de esta solución inicial eran pobres según los estándares habituales, por lo que se eliminaron los ítems con coeficientes de discriminación bajos y se reanalizó el modelo. Este modelo modificado contó con solo 14 ítems y proporcionó un mejor ajuste ( $\chi^2 = 173,74$ ;  $\text{gl} = 135$ ;

$p < 0,001$ ;  $\chi^2/df = 1,28$ ; RMSEA = 0,033; CFI = 0,928; NNFI = 0,918; GFI = 0,934; WRMR = 0,0473). Entre estos ítems restantes, 7 fueron fotografías de varones y los otros 7, de mujeres. Las cargas factoriales para estos ítems oscilaron entre 0,364 y 0,612. La consistencia interna ( $\alpha = 0,689$ ) y la saturación factorial ( $\omega = 0,694$ ) se mantuvieron aceptables. Por último, debido a las modificaciones anteriores se logró aumentar el ajuste general del modelo y se alcanzaron la gran mayoría de los valores de corte estándar (valores de corte: RMSEA < 0,050, CFI > 0,95, NNFI > 0,95, GFI > 0,90 y WRMR < 0,90)<sup>39,40</sup>.

## Discusión

En el presente artículo se analizan algunas propiedades psicométricas del instrumento RMET en su versión infantil. De las características que muestra, se puede destacar que el test estableció correlaciones con instrumentos similares, como es el caso del CAM-C, mientras que no se encontraron correlaciones con medidas de empatía. Para el caso de la consistencia interna, evaluada a través de los coeficientes alfa y omega, evidenció resultados aceptables, con unos de los coeficientes más altos encontrados entre las investigaciones con muestra infantil. A pesar de eso, dichos resultados evidencian que, aparte del uso de índices psicométricos más adecuados para variables categóricas como la omega de McDonald, los resultados del instrumento continúan siendo pobres o aceptables. Por último, en relación con validez de constructo, se confirmó una estructura factorial unidimensional a partir de un análisis confirmatorio.

En cuanto a los ítems, la gran mayoría mostraron un buen poder de discriminación y fueron fáciles de entender por los participantes del estudio. Sin embargo, algún estudio<sup>17</sup> ha mostrado que una versión reducida del test presenta un mejor ajuste del modelo.

Los estimadores puntuales como la media y la desviación estándar encontrados en la presente investigación son muy similares a los reportados por diferentes estudios con población infantil<sup>11,13,15</sup>, en los que las respuestas correctas están entre 16 y 20 puntos. Estos datos respaldan la idea de que el RMET puede producir resultados similares en muestras con diferentes trasfondos culturales, evaluar la ToM de una manera psicométricamente sólida y medir la ToM en una población con desarrollo típico.

En términos de edad, dado que los individuos de este estudio tenían edades similares, no se esperaba encontrar diferencias importantes en los resultados; en efecto, las diferencias se encontraron únicamente entre los grupos extremos de la muestra (10 frente a 16 años), y los individuos mayores obtuvieron puntajes más altos en la prueba. Este hallazgo concuerda con resultados de estudios anteriores<sup>12,41</sup>.

La literatura sobre los cambios en el desarrollo de las habilidades de la ToM en la adolescencia es escasa, pero la evidencia indica que los cambios cuantitativos aumentan durante la transición de la niñez tardía a la adolescencia media<sup>42,43</sup>. Asimismo existen algunos factores que pueden influir en el desarrollo de la prueba, como los efectos de las diferencias individuales en la maduración cerebral<sup>44</sup> o incluso la sensibilidad de la prueba para detectar cambios provocados por la maduración<sup>45</sup>.

El hecho de que los grupos de edad menos parecidos (10 frente a 16 años) fueran los únicos que mostraron diferencias en los resultados podría evidenciar que las habilidades relacionadas con la cognición social permanecen estables durante un periodo considerable de la pubertad y comienzan a aumentar en la adolescencia tardía.

La contrastación de datos entre el presente estudio y anteriores investigaciones en términos de edad se torna difícil, puesto que algunos estudios previos usan intervalos de edad más cortos, algunos toman el grado escolar como variable de comparación y otros agrupan a individuos con diferentes edades cronológicas en el mismo conjunto.

Del mismo modo, en el presente estudio no se hallan diferencias por sexo en el desempeño en la prueba. Esto se contrapone a investigaciones previas como, por ejemplo, los casos de los metanálisis de Mcclure<sup>46</sup> en niños y adolescentes o Hoffmann et al.<sup>47</sup> con adultos, en los que se ha evidenciado una ventaja estadísticamente significativa de las mujeres (si bien no en el test de las miradas) en el reconocimiento facial de emociones, o en el caso específico del test de las miradas en la versión para adultos, en el que se pueden ver diferencias también a favor de las mujeres, como lo demuestra el metanálisis de Kirkland et al.<sup>48</sup>, en el cual se analizan los resultados de varios estudios con población neurotípica de adultos de todo el mundo.

Vellante et al.<sup>18</sup> y Olderbak et al.<sup>17</sup> muestran en sus respectivas revisiones que el test de las miradas presenta importante evidencia sobre la validez convergente con medidas de empatía; esto contrasta con los resultados de algunas investigaciones<sup>11,49</sup> en las que se observa que el test de las miradas no tiene correlaciones con el IRI, como también es el caso del presente estudio.

Tales resultados concuerdan con los reportados en la revisión de Jankowiask-Siuda et al.<sup>50</sup>, que muestra una serie de validaciones en el mundo (Etiopía, Japón, Hungría, Turquía y Suiza) en las que no se encuentra validez convergente con el cociente de empatía (EQ) en ninguna de las versiones mencionadas.

A este respecto, Prevost et al.<sup>51</sup> declaran que, en mayor medida, las diferencias de los resultados entre las pruebas de empatía y RMET, en los diversos estudios reportados, se deben a las diferencias de diseño, contenido y otros puntos importantes.

En cuanto al desempeño presentado por los ítems, se evidenció que, de un total de 28 ítems que tenía el test de las miradas, 5 (1, 2, 3, 25 y 26) presentaron puntajes por debajo del 50% de respuesta correcta. Los ítems 1, 3, 25 y 26 coinciden en su pobre desempeño con los resultados de diferentes estudios de validación de la versión infantil del RMET en el mundo<sup>13,15</sup>.

En aras de dilucidar este problema, Hallerbäck et al.<sup>52</sup> aplicaron la versión sueca del test de las miradas para niños a una muestra de adultos (que deberían tener mayor desarrollo de habilidades para el reconocimiento de emociones y la atribución de estados mentales de terceros) para observar si el comportamiento de estos ítems se debía a su dificultad o a otro tipo de variables. Como resultado, para los ítems 3, 25 y 26 fueron seleccionadas las respuestas correctas más del 50% de las veces, lo que indica que dichos ítems, en efecto, presentan dificultad para la franja de edad de niños y

jóvenes; sin embargo, evalúan adecuadamente el constructo ToM.

Asimismo y en concordancia con Vogindroukas et al.<sup>14</sup>, quienes aplicaron también el RMET versión niños a una muestra de adultos griegos, el ítem 1 mostró un bajo porcentaje de aciertos; en el caso de Suecia, el 35,4% y en el caso de Grecia, el 17,3%, lo que indica que el problema del ítem no es su dificultad para la franja de edad de jóvenes y niños, como sí es el caso de los ítems 3, 25 y 26 ya mencionados, sino otra clase de problema; por ejemplo, en el estudio sueco se indica que un factor que tener en cuenta es la cara de la actriz, que en la foto aparece muy poco iluminada.

No obstante, el desempeño de los ítems también puede explicarse por diferencias culturales, como se ha encontrado previamente<sup>53-58</sup>. De acuerdo con el actual cuerpo de investigación, existe una mayor facilidad de reconocimiento emocional de las expresiones faciales entre miembros de la misma cultura que con personas externas a ella. Asimismo las investigaciones indican que el conjunto de características del entorno en el que se desarrolla el individuo, como la frecuencia de exposición a emociones específicas y la familiaridad con ellas, es una ventaja para el reconocimiento de estados mentales en otros. A modo de ejemplo, la presente investigación muestra más similitudes con la investigación española<sup>14</sup> que con las desarrolladas en otros países, debido a las similitudes culturales y lingüísticas entre Colombia y España.

### Recomendaciones

Artículos previos han señalado que el test de las miradas puede obtener un mejor desempeño psicométrico si se usa una versión breve o reducida. Asimismo, factores como la calidad de la traducción y el formato de aplicación deben tenerse en cuenta en futuras investigaciones. También se recomienda emplear diferentes enfoques y técnicas estadísticas para obtener una comprensión más profunda del desempeño psicométrico del RMET en poblaciones tanto clínicas como no clínicas. Además, se debe probar la estructura factorial hallada en el presente estudio en muestras similares.

### Conclusiones

El RMET posee una consistencia interna y una estructura interna aceptables que se puede mejorar por medio de la adopción de una forma breve. Se deben determinar sus propiedades psicométricas en muestras con desarrollo típico y muestras clínicas de otras regiones del país.

Además, la prueba muestra una moderada validez convergente con otros instrumentos que evalúan cognición social y reconocimiento de emociones, mientras que en el caso de la empatía no parece mostrar correlaciones estadísticamente significativas.

### Conflicto de intereses

Los autores declaran no tener ningún conflicto de interés.

### BIBLIOGRAFÍA

- Adolph R. The neurobiology of social cognition. *Curr Opin Neurobiol.* 2001;11:231-9.
- Weightman M, Air TM, Baune T. A review of the role of social cognition in major depressive disorder. *Front Psychiatry.* 2014;5:1-12.
- Frith C, Frith U. Social Cognition in Humans. *Curr Biol.* 2007;17:724-32.
- Frith DC, Frith U. Implicit and explicit processing in social cognition. *Neuron.* 2008;60:503-10.
- Bertone M, Díaz-Granados E, Vallejos M, Muniello J. Differences in social cognition between male prisoners with antisocial personality or psychotic disorder. *Int J Psychol Res.* 2017;10:16-25.
- Ebert A, Brüne M. Oxytocin and Social Cognition. En: Hurlmann R, Grinevich V, editores. *Behavioral Pharmacology of Neuropeptides: Oxytocin.* New York: Springer; 2017. p. 375-8.
- Shamay-Tsoory SG, Tomer R, Berger B, Goldsher D, Aharon-Peretz J. Impaired "affective theory of mind" is associated with right ventromedial prefrontal damage. *Cogn Behav Neurol.* 2005;18:55-67.
- Sebastian CL, Fontaine M, Bird G, Blakemore SJ, De Brito SA, McCrory E, et al. Neural processing associated with cognitive and affective Theory of Mind in adolescents and adults. *Soc Cogn Affect Neurosci.* 2011;7:53-63.
- Brown M, Thibodeau R, Pierucci J, Gilpin A. Supporting the development of empathy: The role of theory of mind and fantasy orientation. *Soc Dev.* 2017;26:951-64.
- Beaudoin C, Leblanc E, Gangner C, Beauchamp M. Systematic Review and Inventory of Theory of Mind Measures for Young Children. *Front Psychol.* 2020;10.
- Müller C, Gmünder L. An evaluation of the "reading the mind in the eyes test" seventh to ninth graders. *J Intellect Disabil Res.* 2013;7:34-44.
- Van der Meulen V, Roerig S, Ruyter D, Van Lier P, Krabbendam L. A Comparison of Children's Ability to Read Children's and Adults' Mental States in an Adaptation of the Reading the Mind in the Eyes Task. *Front Psychol.* 2017;7:1-15.
- Rueda P, Cabello R, Fernández-Berrocá P. Preliminary validation of Spanish "Eyes Test-Child version". *Ansiedad y Estres.* 2013;19:173-84.
- Vogindroukas I, Chelas EN, Petridis N. Reading the Mind in the Eyes Test (Children's Version): A Comparison Study between Children with Typical Development. Children with High-Functioning Autism and Typically Developed Adults. *Folia Phoniatri Logop.* 2014;66:18-24.
- Girli A. Psychometric Properties of the Turkish Child and Adult Form of "Reading the Mind in the Eyes Test". *Psychology.* 2014;5:1321-37.
- Đorđević J, Živanović M, Pavlović A, Mihajlović G, Stašević I, Pavlović D. Psychometric evaluation and validation of the Serbian version of "Reading the Mind in the eyes" test. *Psychologija.* 2017;4:483-502.
- Olderbak S, Wilhelm O, Olaru G, Geiger M, Brennenman M, Roberts R. A psychometric analysis of the reading the mind in the eyes test: toward a brief form for research and applied settings. *Front Psychol.* 2015;5:1-11.
- Vellante M, Baron-Cohen S, Melis M, Marrone M, Petretto DR, Masala C, et al. The "Reading the Mind in the Eyes" test: Systematic review of psychometric properties and a validation study in Italy. *Cogn Neuropsychiatry.* 2013;18:326-54.
- Baron-Cohen S, Wheelwright S, Spong A, Scahill V, Lawson J. Are Intuitive Physics and Intuitive Psychology Independent? *J Autism Dev Disord.* 2001;5:47-78.

20. Pineda A, Figueroa G, Puentes P. Retrospectiva y prospectiva de la teoría de la mente: Avances de investigaciones en neurociencias. *Psicogente*. 2012;15:178-97.
21. Peñuelas-Calvo I, Sareen A, Sevilla-Llewellyn-Jones J, Fernández-Berrocal P. The "Reading the Mind in the Eyes" Test in Autism-Spectrum Disorders Comparison with Healthy Controls: A Systematic Review and Meta-analysis. *J Autism Dev Disord*. 2019;49:1048-61.
22. Golan O, Sinai-Gavrilov Y, Baron-Cohen S. The Cambridge Mindreading Face-Voice Battery for Children (CAM-C): complex emotion recognition in children with and without autism spectrum conditions. *Mol Autism*. 2015;6:1-9.
23. Cuff BMP, Brown SJ, Taylor L, Howat DJ. Empathy: A Review of the Concept. *Emot Rev*. 2016;8:1-10.
24. Mathôt S, Schreij D, Theeuwes J, OpenSesame. An open-source, graphical experiment builder for the social sciences. *Behav Res*. 2012;44:314-24.
25. Crocker L, Algina J. Introduction to Classical and Modern Test Theory. Masson Cengage Learning. 2008.
26. Bentler PM. Alpha Dimension-Free, and Model-Based Internal Consistency Reliability. *Psychometrika*. 2009;74:137-43.
27. McNeish D. Thanks Coefficient Alpha. We'll Take It From Here. *Psychol Methods*. 2018;23:412-33.
28. Trizano-Hermosilla I, Alvarado JM. Best Alternatives to Cronbach's Alpha Reliability in Realistic Conditions: Congeneric and Asymmetrical Measurements. *Front Psychol*. 2016;7:769.
29. Lee L, Harkness KL, Sabbagh MA. Mental state decoding abilities in clinical depression. *J Affect Disord*. 2005;86:247-58.
30. Timmerman ME, Lorenzo-Seva U. Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychol Methods*. 2011;16:209-20.
31. Li CH. Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behav Res Methods*. 2016;48:936-49.
32. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Modeling*. 1999;6:1-55.
33. JASP Team. JASP (Version 0.14)[Computer software]. 2020.
34. Ferrando PJ, Lorenzo-Seva U. Program FACTOR at 10: origins, development and future directions. *Psicothema*. 2017;29:236-41.
35. Lüdtke D. sjPlot: Data Visualization for Statistics in Social Science.
36. R., Core, Team., R., A., language and environment for statistical, computing., Vienna, Austria. 2020.
37. Malgady R. How Skewed Are Psychological Data? A Standardized Index of Effect Size. *J Gen Psychol*. 2007;134:355-9.
38. Price LR. Psychometric methods: theory into practice. New York: The Guilford Press;. 2017.
39. Kline RB. Principles and Practice of Structural Equation Modeling. New York: The Guildford Press;. 2015.
40. Tabachnick BG, Fidell LS. Using Multivariate Statistics. Boston: Pearson Education;. 2007.
41. Voracek M, Dressler S. Lack of correlation between digit ratio (2D:4D) and Baron-Cohen's "Reading the Mind in the Eyes" test, empathy, systemising, and autism-spectrum quotients in a general population sample. *Pers Individ Differ*. 2006;41:1481-91.
42. Brizio A, Gabbatore I, Tirassa M, Bosco FM. "No more a child, not yet an adult": studying social cognition in adolescence. *Front Psychol*. 2015;6:1011.
43. Vetter NC, Leipold K, Kliegel M, Phillips LH, Altgassen M. Ongoing development of social cognition in adolescence. *Child Neuropsychol*. 2013;19:615-29.
44. Kilford EJ, Garrett E, Blakemore SJ. The Development of Social Cognition in Adolescence: An Integrated Perspective. *Neurosci Biobehav Rev*. 2016;70:106-20.
45. Pino MC, Masedu F, Vagnetti R, Attanasio M, Di Giovanni C, Valenti M, et al. Validity of Social Cognition Measures in the Clinical Services for Autism Spectrum Disorder. *Front Psychol*. 2020;11:4.
46. McClure E. A Meta-Analytic Review of Sex Differences in Facial Expression Processing and Their Development in Infants, Children, and Adolescents. *Psychol Bull*. 2000;126:424-53.
47. Hoffmann H, Kessler H, Eppel T, Rukavina S, Traue C. Expression intensity, gender and facial emotion recognition: Women recognize only subtle facial emotions better than men. *Acta Psychologica*. 2010;135:278-83.
48. Kirkland R, Peterson E, Baker C, Miller S, Pulos S. Meta-analysis Reveals Adult Female Superiority in "Reading the Mind in the Eyes Test. *North Am J Psychology*. 2013;15:121-46.
49. Lee HN, Nam G, Hur JW. Development and validation of the Korean version of the Reading the Mind in the Eyes Test. *PLoS One*. 2020;15:1-13.
50. Jankowiask-Siuda K, Baron-Cohen S, Bialaszek W, Dopierala A, Kozłowska A, Rymarczyk K. Psychometric evaluation of the "reading the mind in the eyes" test with samples of different ages from a Polish population. *Stud Psychol*. 2016;58:18-32.
51. Prevost M, Carrier ME, Chowne G, Zelkowitz P, Joseph L, Gold I. The Reading the Mind in the Eyes test: validation of a French version and exploration of cultural variations in a multi-ethnic city. *Cogn Neuropsychiatry*. 2013;19:1-16.
52. Hallerbäck M, Lugnegard T, Hjärthag F. The Reading the Mind in the Eyes Test: Test-retest reliability of a Swedish version. *Cogn Neuropsychiatry*. 2009;14:127-43.
53. Kobayashi C, Glover G, Temple E. Cultural and Linguistic Influence on Neural Bases of "Theory of Mind": an fMRI Study with Japanese bilinguals. *Brain Lang*. 2006;98:210-20.
54. Kobayashi F, Temple E. Cultural Effects on the Neural Basis of Theory of Mind. *Prog Brain Res*. 2009;178:213-23.
55. Adams R, Rule N, Franklin N, Wang E, Stevenson M, Yoshikawa S, et al. Cross-cultural Reading the Mind in the Eyes: An fMRI Investigation. *J Cogn Neurosci*. 2010;22:97-108.
56. Paulman S, Uskul A. Cross-cultural emotional prosody recognition: evidence from Chinese and British listeners. *Cogn Emot*. 2014;28:230-44.
57. Bjornsdottir R, Rule N. On the relationship between acculturation and intercultural understanding: Insight from the Reading the Mind in the Eyes test. *Int J Intercult Relat*. 2016;52:39-48.
58. Bradford E, Jentsch I, Gómez J, Chen Y, Zhang D, Su Y. Cross-Cultural Differences in Adult Theory of Mind Abilities: A Comparison of Native-English Speakers and Native Chinese Speakers on the Self/Other Differentiation Task. *Q J Exp Psychol*. 2018;71:1-26.