

## Baremación de las Matrices Progresivas de Raven (Escala Coloreada) en el Distrito San Roque de Asunción

Development of Norms for Raven's Coloured Progressive Matrices in San Roque District (Asunción)

Guadalupe Escobar y Eduardo Cáceres  
Universidad Católica "Nuestra Señora de la Asunción"

CDID "Centro de Documentación, Investigación y Difusión de Psicología Científica"<sup>1</sup>

Recibido: 15/junio/2024

Aceptado: 5/Octubre/2024

### Resumen

**Introducción:** El Test de Matrices Progresivas de Raven es un instrumento psicométrico normativo utilizado para la medición de la inteligencia general. **Objetivo:** La presente investigación constituye una baremación de la Escala Coloreada del instrumento en niños de 8 a 11 años matriculados en instituciones educativas públicas del Distrito San Roque de Asunción, Paraguay. **Método:** Se utilizó un muestreo bietápico semi probabilístico superior, con una muestra de 661 niños (122 de 8 años, 222 de 9 años, 207 de 10 años y 110 de 11 años) pertenecientes a 14 instituciones educativas públicas. El instrumento, en formato cuadernillo, fue administrado en grupos de hasta 11 niños, sin límites de tiempo. Para el análisis se realizaron pruebas estadísticas descriptivas e inferenciales, calculando: distribución de frecuencias, media, desviación estándar y puntajes mínimos y máximos de cada grupo etario y de sus sub-grupos de sexo; también se realizaron la prueba de Shapiro-Wilk de bondad de ajuste y las pruebas de Kruskal-Wallis y de U-Mann Whitney para las comparaciones de puntajes. **Resultados y Conclusión:** Se obtuvieron datos normativos relevantes para la población seleccionada. No se encontraron diferencias significativas en los puntajes entre hombres y mujeres en la mayoría de los grupos etarios. Se encontraron diferencias significativas entre los puntajes de los niños de 8 y 10 años y de 8 y 11 años, pero no entre los demás pares de grupos etarios. El puntaje promedio de cada grupo etario fue de 25 (8 años), 26.6 (9 años), 27.5 (10 años) y 27.3 (11 años).

*Palabras clave:* baremación, datos normativos, Test de Matrices Progresivas de Raven, Escala Coloreada, estudio psicométrico.

<sup>1</sup> Correspondencia remitir a: [guada.escobar00@gmail.com](mailto:guada.escobar00@gmail.com)

<sup>2</sup> Correspondencia remitir a: [revistacientificaureka@gmail.com](mailto:revistacientificaureka@gmail.com) o [normacopparipy@gmail.com](mailto:normacopparipy@gmail.com) "Centro de Documentación, Investigación y Difusión de Psicología Científica", de Asunción-Paraguay.

## Abstract

**Introduction:** The Raven's Progressive Matrices test is a norm-referenced psychometric assessment used to measure general intelligence. **Objective:** The goal of this study was to obtain normative data for Raven's Coloured Progressive Matrices for children aged 8 to 11 from state schools in San Roque District (Asunción, Paraguay). **Method:** Two-stage upper semiprobability sampling was used. The final sample consisted of 661 students from 14 different state schools (122 children aged 8, 222 aged 9, 207 aged 10 and 110 aged 11). The test, in the booklet form, was applied in groups of up to 11 children, without time constraints. For the analysis, both descriptive and inferential statistics were utilized, calculating: the frequency distribution, the mean, the standard deviation and the minimum and maximum scores for each age group, as well as for their sex sub-groups; the Shapiro-Wilk test for normality and the Kruskal-Wallis and U-Mann Whitney tests for comparisons of scores were also utilized. **Results and Conclusion:** Relevant normative data was obtained for the selected population. The findings show no significant differences between the scores of males and females in most age groups tested. There were significant differences between the scores of children of 8 and 10 years, as between the scores of children of 8 and 11 years, but no significant differences were found in other age group comparisons. The mean score of each age group was 25 (8 years), 26.6 (9 years), 27.5 (10 years) and 27.3 (11 years).

*Key words:* norms, normative data, Raven's Progressive Matrices, Coloured Progressive Matrices, psychometric study.

Los instrumentos psicométricos son herramientas fundamentales en todos los procesos de evaluación psicológica, por lo cual su correcta utilización es sumamente importante, siendo uno de los aspectos que debe tenerse en cuenta el de la interpretación de sus resultados, especialmente cuando los mismos se obtienen mediante la comparación con puntajes normativos de una población. Un instrumento cuyos resultados se interpretan de esta manera es el Test de Matrices Progresivas de Raven (Raven et al., 1993), el cual constituye uno de los principales instrumentos utilizados en la práctica psicológica para la medición de la inteligencia general y el razonamiento abstracto (Oakland et al., 2015). Este instrumento fue desarrollado por John C. Raven en 1938, con el objetivo de “medir la capacidad intelectual... para comparar formas y razonar por analogía, con independencia de los conocimientos adquiridos” (Raven et al., 1991/1993, p. 9). Se fundamenta en la teoría ecléctica de los dos factores de la inteligencia de Charles Spearman, así como en las leyes no genéticas formuladas por el mismo autor.

Al ser un instrumento normativo, las puntuaciones obtenidas a partir de la administración del Test de Raven, llamadas puntuaciones brutas, no son directamente interpretables puesto que no aportan información por sí mismas, sino que adquieren significado a través de la comparación con las puntuaciones obtenidas por un grupo de referencia determinado (González, 2007). Esta comparación se realiza a través de las normas, las cuales describen la distribución de las puntuaciones obtenidas por un grupo en el instrumento, identificando la actuación media o normal (estadísticamente) del comportamiento de la población determinada en la variable evaluada (González, 2007).

Las normas para los instrumentos psicométricos se obtienen a través de un proceso de tipificación y se encuentran resumidas y expresadas en los baremos de un test. Los baremos son tablas que resumen, condensan y organizan los datos obtenidos en patrones significativos, asignando a cada posible puntuación directa un valor numérico en una determinada escala que informa acerca de la posición que la persona ocupa, en cuanto a la variable evaluada, en relación al grupo en base al cuál se establecieron esas normas (Gregory, 2012).

Las normas de los instrumentos no sólo proporcionan información acerca de la posición de un individuo en referencia al grupo, sino que también proporcionan información a un nivel más amplio, permitiendo adquirir un entendimiento de la manera en que los distintos grupos de una población determinada difieren en ciertas variables, así como también del grado en que estas diferencias se relacionan a las características de los grupos y poblaciones (Angoff, 1984).

Las normas de los instrumentos no tienen una validez ilimitada. En primer lugar, porque describen el rendimiento promedio de una población determinada: la utilizada para su obtención; así, pueden no ser adecuadas como puntos de comparación para poblaciones de características distintas. En segundo lugar, no son permanentes, dado que los valores normativos pueden variar a lo largo del tiempo debido a diversos factores, requiriéndose así de actualizaciones constantes (Cerdá, 1978).

Cuando se utilizan normas inadecuadas, ya sea porque el grupo normativo no es análogo al del individuo evaluado o porque las normas están desactualizadas, las variaciones pueden ser significativas, resultando en interpretaciones erróneas de los resultados que pueden incluso derivar en diagnósticos erróneos de los individuos evaluados (Anastasi, 1976; Daugherty et al., 2016).

Las normas del Test de Raven, como las de todos los instrumentos psicométricos, no son absolutas, permanentes ni aplicables a todas las poblaciones, motivo por el cual para una interpretación adecuada de sus resultados debe contarse con normas locales y también actualizadas. En el manual del Test de Matrices Progresivas de Raven, los mismos autores hacen referencia a investigaciones que indican un aumento intergeneracional de los puntajes obtenidos en el instrumento, estableciendo que la capacidad educativa se ve incrementada a razón de más o menos una desviación estándar por generación en numerosas partes del mundo, existiendo diversas hipótesis que explican el fenómeno encontrado (Raven et al., 1993).

En la misma línea, a pesar de que el Test de Raven es menos sensible a las diferencias culturales que otros instrumentos (Kaplan y Saccuzzo, 2018), se han encontrado diferencias en los puntajes de individuos de diferentes culturas (Kihn, 2005; Lozano-Ruiz et al., 2021; Raven, 2000; Raven et al., 1998), reforzando la importancia de utilizar normas pertinentes y actualizadas para interpretar los resultados de este instrumento.

En Paraguay, las normas más actualizadas del Test de Matrices Progresivas de Raven corresponden a la Escala General (Duarte, 2017; Mongelós et al., 2020). En cuanto a la Escala Especial Coloreada, objeto del presente trabajo, se cuenta con una baremación nacional realizada por el Ministerio de Educación y Ciencias; sin embargo, ésta fue realizada hace más de veinte años y no se encuentra debidamente documentada, siendo el acceso a ella sumamente difícil. Atendiendo a esta carencia, el presente trabajo se realiza con el objetivo de elaborar baremos del Test de Matrices Progresivas de Raven (Escala Coloreada) para alumnos de 8 a 11 años de edad de escuelas públicas del distrito San Roque de Asunción.

## **Método**

### **Diseño**

Es una investigación cuantitativa no experimental. El diseño utilizado es de tipo instrumental descriptivo (Hernández-Sampieri et al., 2014; Montero y León, 2005).

### **Participantes**

La población seleccionada fue la de estudiantes de 8 a 11 años de edad, matriculados en escuelas públicas del distrito San Roque de la ciudad de Asunción. Para la determinación de la muestra se realizó un muestreo bietápico de tipo semi probabilístico superior (Cochran et al., 1954), tomando a las escuelas y a sus alumnos como unidades primarias y secundarias de muestreo, respectivamente, y calculándose los tamaños muestrales con un nivel de confianza del 95%.

El muestreo es semi probabilístico superior dado que las escuelas fueron seleccionadas mediante una técnica probabilística de muestreo aleatorio simple, sin embargo, los estudiantes fueron seleccionados mediante un muestreo no-probabilístico de conveniencia (en base a la asistencia a las escuelas en los días en que se realizaron las sesiones de administración, y previo consentimiento informado de los padres). Esta técnica de muestreo fue seleccionada teniendo en cuenta las limitaciones administrativas y de recursos del proyecto y las consideraciones de Lindquist (1930) y Lord (1957) respecto a la importancia de una correcta selección de las escuelas como unidades primarias de muestreo para la obtención de normas a nivel escolar. Para asegurar que todas las escuelas se encontrasen suficientemente representadas, se administró el test a un mínimo del 35% de los alumnos de cada grado en cada escuela.

El instrumento fue aplicado a 236 alumnos de 3° grado, 231 alumnos de 4° grado y 245 alumnos de 5° grado. Tras la corrección de la prueba administrada, fueron descartados 51 alumnos por no cumplir el criterio de edad y/o por posibles errores de administración del instrumento reflejados en un valor anormal de discrepancia con los puntajes esperados, acorde a lo indicado en el manual del instrumento.

La muestra final se compone de 661 alumnos matriculados en 14 instituciones educativas oficiales del Distrito San Roque de la ciudad de Asunción, de los cuales 225 pertenecen al 3° grado, 215 al 4° grado y 221 al 5° grado. En la Tabla 1 se presenta la distribución final de la muestra, de acuerdo a las edades y según el sexo de los participantes.

**Tabla 1**

*Distribución de la muestra*

Edad	n		
	Mujeres	Hombres	Total
8 años	68	54	122
9 años	113	109	222
10 años	109	98	207
11 años	53	57	110
<b>Total</b>			<b>661</b>

**Instrumentos y Materiales**

Se utilizó el Test de Matrices Progresivas de Raven, Escala Coloreada (1993), formato cuadernillo. Este instrumento consta de 36 reactivos en forma de matrices lacunarias incompletas con formato de respuesta de selección múltiple, distribuidos en 3 series de dificultad creciente (A, Ab y B). La administración puede realizarse de manera colectiva o individual, sin límite de tiempo. Los resultados obtenidos proporcionan información acerca de la capacidad de razonamiento por analogías del evaluado. En adición, el instrumento permite calcular las discrepancias en la composición del puntaje del sujeto y categorizarlas como normales/anormales. Este dato se obtiene comparando el puntaje obtenido por un individuo determinado en cada serie con la cifra normalmente esperada en base a su puntaje total.

Cuando hay una desviación de más de 2 puntos respecto a lo esperado en alguna de las series, el puntaje total obtenido no puede aceptarse en su valor nominal como medición consistente de la capacidad general para la actividad intelectual del individuo. Este dato permite eliminar del grupo normativo casos no válidos que podrían afectar la distribución final de la muestra (Raven et al., 1993).

En cuanto a la validez, el manual reporta una alta saturación del factor “g”, con algún grado de participación del factor viso-espacial. Las Matrices Progresivas de Raven evalúan considerablemente el factor de procesamiento simultáneo, con saturaciones típicas entre .75 y .85. Los estudios sobre la confiabilidad del instrumento han confirmado que es sumamente satisfactoria, con valores superiores a .80 en la mayoría de los casos, tanto mediante la técnica de división de mitades como mediante el método de test-retest (Raven et al., 1993).

## **Procedimiento**

Se realizaron reuniones con los directores de las instituciones educativas seleccionadas para presentar el proyecto y obtener el consentimiento de participación. Tras la obtención del consentimiento informado de los padres, la aplicación de la prueba se realizó en salones y aulas de cada institución, con ciertos requisitos de iluminación y espacio y de superficies de escritura acordes a los materiales. Las sesiones de administración fueron realizadas en agosto de 2023, en grupos de hasta 11 alumnos. Cada alumno recibió un cuadernillo de matrices y una hoja de respuestas, y se siguieron las instrucciones establecidas en el manual del instrumento para la aplicación.

Los datos registrados de cada participante fueron: nombre completo, fecha de nacimiento, fecha de administración, edad, institución educativa, grado escolar, hora de inicio y de fin, y duración. En los casos en los que los datos se encontraban incompletos, los mismos fueron obtenidos a través de los encargados correspondientes en cada institución, siendo estos datos principalmente los de fecha de nacimiento y de grado. Tras la corrección del instrumento se descartaron 51 casos de la muestra por encontrarse fuera del rango de edad seleccionado y/o por presentar valores de discrepancia anormal.

Para el análisis de datos se utilizó el Paquete Estadístico de Ciencias Sociales (SPSS, Versión 27.0). Se realizaron pruebas estadísticas descriptivas e inferenciales, descritas a continuación. En primer lugar, se calculó la distribución de frecuencias de los puntajes en cada grupo etario, así como la media, la desviación estándar, el mínimo y el máximo. Estos cálculos se realizaron también para los sub-grupos de sexo en cada edad seleccionada. En cuanto a los análisis inferenciales, se realizó la prueba de Shapiro-Wilk de bondad de ajuste para determinar la normalidad de la distribución de los puntajes en cada edad.

En función de los resultados obtenidos, se utilizó la prueba de Kruskal-Wallis para comparar los puntajes obtenidos, según los criterios de edad y de grado. Se realizaron pruebas post-hoc para analizar las diferencias existentes. Posteriormente, se utilizó la prueba de U-Mann Whitney para la comparación de los puntajes obtenidos de acuerdo al sexo de los participantes, con la finalidad de determinar la necesidad de baremos diferenciados por sexo. Finalmente, se procedió a la baremación, calculando los valores percentilares correspondientes a cada puntuación en cada grupo etario.

Finalmente, se entregaron informes grupales a las instituciones participantes, describiendo los resultados obtenidos por cada individuo de la institución participante en términos de percentiles y de rangos correspondientes.

## **Resultados**

Con la finalidad de obtener las normas correspondientes a los grupos de edad, se calcularon los siguientes índices estadísticos de media, desviación estándar, mínimo y máximo de cada grupo etario, distinguiendo estos datos en cada serie de matrices que componen el instrumento, así como en relación al puntaje total en el mismo (Tabla 2).

**Tabla 2**

*Descriptivos para el puntaje del Test de Raven según edad*

Edad	Serie	Media	DS	Mín.	Máx.
8 años	A	9.24	1.51	5	12
	Ab	8.79	2.46	2	12
	B	6.98	2.73	1	12
	Total	25.0	6.00	10	35
9 años	A	9.67	1.37	4	12
	Ab	9.43	2.13	3	12
	B	7.54	2.67	2	12
	Total	26.6	5.40	9	35
10 años	A	9.75	1.55	3	12
	Ab	9.81	2.12	1	12
	B	7.89	2.64	1	12
	Total	27.5	5.52	8	36
11 años	A	10.0	1.39	6	12
	Ab	9.61	2.14	3	12
	B	7.70	2.74	2	12
	Total	27.3	5.47	12	36

En la mayoría de los grupos etarios el puntaje promedio obtenido por los alumnos en el instrumento se reduce de manera progresiva en las 3 series que componen el instrumento, respondiendo al ordenamiento de los reactivos en series de dificultad creciente. En la misma línea, la variabilidad de los puntajes tiende a ser menor en la Serie A, en la cual la mayoría de los niños obtienen puntajes cercanos, aumentando progresivamente en las siguientes series.

Posteriormente se procedió a la comparación de la media de los puntajes obtenidos por las submuestras. Se realizó primeramente una prueba de bondad de ajuste para determinar la normalidad de la distribución de los puntajes obtenidos en la prueba por cada submuestra (Tabla 3), encontrándose distribuciones no normales para todos los grupos seleccionados.

**Tabla 3**

*Pruebas de Shapiro-Wilk (Normal) para Test de Raven*

Variable	W	p-valor	Ajuste
Puntaje (8 años)	.957	.000	No
Puntaje (9 años)	.948	.000	No
Puntaje (10 años)	.901	.000	No
Puntaje (11 años)	.955	.000	No
Puntaje (General)	.943	.000	No

Al no cumplirse los criterios para pruebas paramétricas, las comparaciones se realizaron utilizando la prueba de Kruskal-Wallis. Se comparó el puntaje obtenido en el Test de Matrices Progresivas de Raven, tanto según grado como según edad, encontrándose diferencias significativas en ambas comparaciones (p-valor = .001 en el Factor de Edad; p-valor = .000 en el Factor de Grado).

Se realizaron pruebas post-hoc para determinar los pares de edades y de grados que presentan diferencias significativas en los puntajes (Tabla 4). Se encontraron diferencias significativas en los puntajes totales de los niños de 8 y 10 años y de 8 y 11 años, pero no entre los demás pares de grupos etarios.

Al realizar la comparación por grados, se encontraron diferencias estadísticamente significativas entre los puntajes de los alumnos de 3° y 4° grado y entre los de 3° y 5° grado.

Estos resultados son coherentes con lo postulado en el manual del instrumento, en el que se establece que los individuos presentan un progreso sistemático en sus puntajes del test a medida que hay un aumento de la edad (Raven et al., 1993).

Este aumento progresivo se observa igualmente en gran parte de los estudios de baremación del Test de Matrices Progresivas de Raven (Escala Coloreada). Sin embargo, se destaca que en la población de la presente investigación este progreso no resulta perceptible entre todas las edades consecutivas, sino sólo en algunos pares etarios.

**Tabla 4**

*DSCF (post-hoc) para puntaje total del Test de Raven, según edad (años cumplidos) y grado*

<b>Criterio de comparación</b>	<b>Grupo 1</b>	<b>Grupo 2</b>	<b>p-valor</b>	<b>Dif. (5%)</b>
Según edad	8 años	9 años	.080	No
	8 años	10 años	.000	Sí
	8 años	11 años	.020	Sí
	9 años	10 años	.258	No
	9 años	11 años	.723	No
	10 años	11 años	.980	No
Según cursado	3° grado	4° grado	.000	Sí
	3° grado	5° grado	.000	Sí
	4° grado	5° grado	.996	No

Posteriormente, con la finalidad de determinar la necesidad de dividir los baremos de acuerdo al sexo, se realizaron comparaciones de los puntajes obtenidos teniendo en cuenta el sexo de los participantes (Tabla 5), utilizando la prueba de U-Mann Whitney.

**Tabla 5***Comparación de puntajes del Test de Raven según sexo (U-Mann Whitney)*

<b>Factor</b>	<b>p-valor</b>	<b>Dif. (5%)</b>	<b>r</b>	<b>Mayor en</b>
Serie A	.469	No	-	-
Serie Ab	.369	No	-	-
Serie B	.364	No	-	-
Puntaje total	.546	No	-	-
Puntaje total (8 años)	.672	No	-	-
Puntaje total (9 años)	.841	No	-	-
Puntaje total (10 años)	.042	Sí	.163	Varones
Puntaje total (11 años)	.095	No	-	-

En general, no se encontraron diferencias significativas en los puntajes obtenidos de manera global por individuos de un sexo en comparación al otro, así como tampoco en los puntajes obtenidos en cada serie. Al realizar comparaciones internas dentro de cada grupo etario, sin embargo, se encontró que las diferencias en los puntajes globales sí son significativas entre los niños de 10 años, con un tamaño de efecto bajo ( $r = .16$ ), siendo los varones los que puntúan más alto a esta edad.

Finalmente, respondiendo al objetivo de la investigación de elaborar baremos del Test de Matrices Progresivas de Raven para la población seleccionada, se realizó la transformación de las puntuaciones brutas, utilizando el sistema de medida de percentiles para ello. Al no encontrarse diferencias significativas entre los individuos según el sexo en la mayoría de los grupos etarios, la baremación se realizó distinguiendo sólo según edades.

En la Tabla 6 se presentan los baremos correspondientes a la población de alumnos de 8 a 11 años matriculados en escuelas públicas del Distrito San Roque de Asunción.

**Tabla 6**

*Percentiles correspondientes a los puntajes del Test de Raven, según edad*

Percentil	Edad			
	8 años	9 años	10 años	11 años
5	14	15	16	16
10	16	19	21	21
15	18	21	22	22
20	19	23	24	24
25	20	24	25	25
30	23	25	25	25
35	24	25	26	26
40	24	26	27	27
45	25	27	28	28
50	26	27	28	28
55	27	28	29	29
60	27	29	30	30
65	28	30	30	30
70	29	30	31	31
75	30	31	32	32
80	31	31	32	32
85	31	32	33	33
90	32	33	33	33
95	33	34	34	35
99	35	35	35	36

## Discusión

En concordancia con investigaciones anteriores (Kazem et al., 2009; Khaleefa y Lynn, 2008; Raven, 2000; Raven et al., 1991/1993), los resultados obtenidos evidencian:

1. La dificultad creciente de las series de reactivos del Test de Matrices Progresivas de Raven (Escala Coloreada), al presentarse una reducción progresiva del puntaje promedio obtenido en cada serie en la mayoría de los grupos etáreos de la población.
2. La existencia de diferencias significativas en el desempeño en la prueba de sujetos de distintas edades, aun cuando las mismas no fueron significativas en las comparaciones entre todos los pares de edades.
3. La ausencia de diferencias significativas en el desempeño en la prueba de sujetos de sexo distinto, descartándose por ello la necesidad de baremos diferenciados para individuos del sexo femenino y del sexo masculino de estas edades en el instrumento.

Estos resultados permiten aceptar la validez de los baremos obtenidos para la población de estudio, al ser los mismos concordantes con los fundamentos del Test de Matrices Progresivas de Raven y con los estudios posteriores de baremación del mismo ya mencionados.

Así, como resultado del presente trabajo de investigación se cuentan con datos normativos para la población de niños de 8 a 11 años pertenecientes a escuelas públicas del Distrito San Roque de la ciudad de Asunción en el Test de Matrices Progresivas de Raven (Escala Coloreada), los cuales constituyen una fuente de información que permitirá la interpretación de los resultados de otros individuos de la población en el instrumento.

Los baremos elaborados utilizan el sistema de medida percentilar, al ser el mismo uno de los más utilizados en los baremos de instrumentos psicométricos (Angoff, 1984; Nitko y Brookhart, 2013) y el recomendado en el manual del instrumento, en el que se establece que “*el método más satisfactorio para interpretar el significado de un puntaje es considerarlo en función de la frecuencia porcentual con la que ese valor aparece entre personas de la misma edad*” (Raven et al., 1993, p. 68). Este sistema fue seleccionado también debido a las ventajas derivadas de su uso, entre las cuáles se destaca la facilidad con la que los resultados obtenidos pueden ser comprendidos por personas ajenas al ámbito de la psicología, especialmente teniendo en cuenta que el Test de Matrices Progresivas de Raven se utiliza frecuentemente en contextos escolares y con profesionales ajenos a la psicología.

La posibilidad de contar con datos normativos no sólo actualizados, sino pertinentes a la población, constituye el aspecto más importante, teniendo en cuenta las características de no-universalidad y no-permanencia de los baremos de todos los tests psicométricos y recordando que, en los casos en los que no se cuentan con normas actualizadas y relevantes para una población, las variaciones pueden ser significativas y conducir a conclusiones erróneas (Anastasi, 1976). La valoración de la capacidad intelectual de un individuo posee ramificaciones críticas en la vida del mismo, especialmente cuando la misma es categorizada como deficiente o inferior.

Estas valoraciones se utilizan para determinar diagnósticos, estimar comportamientos futuros, recomendar servicios de asistencia, realizar adecuaciones curriculares, establecer métodos de intervención y ubicar a niños en programas especializados, entre muchas otras acciones (Shuttleworth-Edwards, 2019). La disponibilidad de normas locales actualizadas para una población del país permitirá la realización de evaluaciones más precisas y equitativas de estos individuos, aumentando en consecuencia las posibilidades de intervenciones apropiadas y adecuadas a sus particularidades y las oportunidades de desarrollar su capacidad al máximo.

En cuanto a las limitaciones, se destaca principalmente el alcance del proyecto. A pesar del gran tamaño de la muestra, los baremos resultantes constituyen baremos locales, cuya validez sólo ha sido comprobada para una población determinada y limitada.

Con esto no se pretende menoscabar la importancia de los baremos locales, sino destacar la aún presente necesidad de contar con datos normativos actualizados para el resto de las poblaciones de nuestro país en el instrumento, así como también con baremos nacionales. Si bien los baremos locales son sumamente importantes tanto en contextos clínicos como escolares, en los que una comparación del rendimiento del niño con personas de su contexto diario permite implementar medidas en el aula ajustadas verdaderamente a las necesidades precisas del niño en ella, la obtención de baremos nacionales que permitan la identificación y el análisis del nivel de razonamiento abstracto general de los niños de todo el país es fundamental para que puedan diseñarse e implementarse políticas educativas nacionales fundamentadas en la evidencia.

Otra limitación se deriva del rango etario seleccionado para el proyecto, el cual no abarca el rango completo de edades para las cuales el instrumento fue diseñado, viéndose excluidas las edades de 5 a 7 años debido a las características de la aplicación de la Forma Tablero del instrumento y considerando las limitaciones de tiempo y de recursos del presente proyecto.

Los resultados obtenidos sugieren la necesidad de continuar realizando proyectos de baremación de instrumentos en poblaciones paraguayas. Aunque el presente proyecto no constituye un proyecto de comparación de los baremos actuales con los utilizados de manera predominante en el país, debido a las limitaciones metodológicas y a las características diferentes de ambos baremos, un análisis exploratorio revela que existen diferencias significativas entre los mismos. Si los baremos utilizados para interpretar los puntajes de los individuos de la muestra hubiesen sido los existentes anteriormente, los diagnósticos alcanzados del nivel de razonamiento abstracto de los individuos hubieran sido más laxos y, en consecuencia, podrían no ser acordes a la realidad de la población.

Con este trabajo se han sentado las bases para la realización de proyectos de baremación similares, ya sean correspondientes a baremaciones locales, de subgrupos o nacionales. A investigadores interesados en el tema, se recomiendan también líneas de investigación adyacentes, como la comparación de baremos de distintos años o de distintas poblaciones, estudios de comprobación del efecto Flynn, o la exploración de las posibles causas de diferencias en los datos normativos de distintas poblaciones, entre otras.

## Referencias

- Anastasi, A. (1976). *Psychological testing* (4.<sup>a</sup> ed.). Macmillan Publishing Co., Inc.
- Angoff, W. H. (1984). *Scales, norms, and equivalent scores*. Educational Testing Service.
- Cerdá, E. (1978). *Psicometría general*. Herder.
- Cochran, W.G., Mosteller, F., y Tukey, J.W. (1954). Principles of Sampling. *Journal of the American Statistical Association*, 49(265), 13-35.  
<https://doi.org/10.1080/01621459.1954.10501212>
- Daugherty, J.C., Puente, A.E., Fasfous, A.F., Hidalgo-Ruzzante, N., y Pérez-García, M. (2016). Diagnostic mistakes of culturally diverse individuals when using North American neuropsychological tests. *Applied Neuropsychology: Adult*, 24(1), 16-22. <http://dx.doi.org/10.1080/23279095.2015.1036992>
- Duarte, M. V. (2017). *Construcción de Baremos del Test de Matrices Progresivas de Raven, Escala General en Estudiantes del Bachillerato de la Ciudad de Asunción*. [Tesis de grado, Universidad Nacional de Asunción].
- González, F. M. (2007). *Instrumentos de evaluación psicológica*. Editorial Ciencias Médicas.
- Gregory, R. J. (2012). *Pruebas psicológicas: Historia, principios y aplicaciones* (1.<sup>a</sup> ed.). Pearson Educación.
- Hernández-Sampieri, R., Fernández, C., y Del Pilar, M. (2014). *Metodología de la investigación* (6.<sup>a</sup> ed.). Mcgraw-Hill.

- Kaplan, R. M., y Saccuzzo, D. P. (2018). *Psychological testing: principles, applications and issues* (9.<sup>a</sup> ed.). Cengage Learning.
- Kazem, A. M., Alzubiadi, A. S., Alkharusi, H. A., Yousif, Y. H., Alsarmi, A. M., Al-Bulushi, S. S., Aljamali, F. A., Al-Mashhdany, S., Al-Busaidi, O. B., Al-Fori, S. M., Al-Bahrani, W. A., y Alshammary, B. M. (2009). A normative study of the Raven's Coloured Progressive Matrices test for Omani children aged 5-11 years. *Jurnal Pendidikan Malaysia (Malaysian Journal of Education)*, 34(1), 37–51.
- Khaleefa, O., y Lynn, R. (2008). Normative Data for Raven's Coloured Progressive Matrices Scale in Yemen. *Psychological Reports*, 103(1), 170–172. <https://doi.org/10.2466/pr0.103.1.170-172>
- Kihn, T. (2005). The appropriateness of the raven's coloured progressive matrices and its existing normative data for Zulu speakers [Tesis de maestría, University of KwaZulu-Natal]. <http://hdl.handle.net/10413/1890>
- Lindquist, E. F. (1930). Factors determining reliability of test norms. *Journal of Educational Psychology*, 21(7), 512–520. <https://doi.org/10.1037/h0070463>
- Lord, F. M. (1957). Test Norms and Sampling Theory. *ETS Research Bulletin Series*, 1957(1), i–48. <https://doi.org/10.1002/j.2333-8504.1957.tb00068.x>
- Lozano-Ruiz, A., Fasfous, A. F., Ibanez-Casas, I., Cruz-Quintana, F., Perez-Garcia, M., y Pérez-Marfil, M. N. (2021). Cultural Bias in Intelligence Assessment Using a Culture-Free Test in Moroccan Children. *Archives of Clinical Neuropsychology*, 1–9. <https://doi.org/10.1093/arclin/acab005>
- Mongelós, P., Ferreira, J., Espínola, J., Colmán, L., Duarte, A., Domínguez, C., Cano, A., Gray, L., y Leguizamón, C. (2020). Test de Raven: Baremo en adolescentes del tercer curso de nivel medio de San Pedro del Ycuamandyyú. *Revista Científica de La Facultad de Filosofía*, 11(2), 115–124. <https://revistascientificas.una.py/index.php/rcff/article/view/2856>
- Montero, I., y León, O. G. (2007). A guide for naming research studies in Psychology. *International Journal of Clinical and Health Psychology*, 7(3), 847–862. Sistema de Información Científica Redalyc. <https://www.redalyc.org/articulo.oa?id=33770318>

- Nitko, A. J., y Brookhart, S. M. (2013). *Educational Assessment of Students* (6.<sup>a</sup> ed.). Pearson Education Limited.
- Oakland, T., Douglas, S., y Kane, H. (2015) Top Ten Standardized Tests Used Internationally With Children and Youth by School Psychologists in 64 Countries: A 24-Year Follow-Up Study. *Journal of Psychoeducational Assessment*, 34(2), 166-176. <https://doi.org/10.1177/0734282915595303>
- Raven, J. (2000). The Raven's Progressive Matrices: Change and Stability over Culture and Time. *Cognitive Psychology*, 41(1), 1–48. <https://doi.org/10.1006/cogp.1999.0735>
- Raven, J., Raven, J. C., y Court, J. H. (1993). *Test de matrices progresivas: Manual*. Editorial Paidós.
- Raven, J., Raven, J. C., y Court, J. H. (1998). *Manual for Raven's progressive matrices and vocabulary scales*. Oxford Psychologists Press.
- Shuttleworth-Edwards, A. B. (2019). Using the WAIS-III to illustrate test norming strategies in multicultural contexts: A demographically stratified sampling design. En S. Laher, A. Fynn, y S. Kramer (Eds.), *Transforming Research Methods in the Social Sciences: Case Studies from South Africa* (pp. 69-83). Wits University Press. <https://doi.org/10.18772/22019032750.10>