

Matrices Progresivas de Raven: efecto Flynn y actualización de baremos

Raven's Progressive Matrices: Flynn effect and updated norms

Lilia Rossi-Casé, Rosa Neer, Susana Lopetegui, Stella Doná, Bruno Biganzoli y Ramiro Garzaniti

Universidad Nacional de La Plata, La Plata, Argentina

Resumen

Este trabajo presenta los resultados obtenidos en la construcción de normas válidas y actualizadas del Test de Matrices Progresivas de Raven, Escala General, para la ciudad de La Plata, Argentina; y realiza una comparación con las normas obtenidas en 1964 y 2000 a fin de observar la tendencia de los puntajes. Se seleccionó una muestra de 1001 estudiantes de ambos sexos y provenientes de distintos establecimientos de educación secundaria, con edades comprendidas entre los 13 y 18 años. El instrumento fue administrado de manera colectiva, en grupos-clase, con presencia del examinador, sin límite de tiempo para su ejecución, de manera tal que permitiera evaluar la capacidad intelectual sin la intervención de la velocidad en la tarea. Los resultados obtenidos se analizaron dividiéndolos en tres grupos por edad (i.e., 13-14 años, 15-16 años, 17-18 años), y se compararon con los baremos de 1964 y 2000. Se observó un leve descenso de los puntajes directos necesarios para alcanzar el rendimiento promedio, respecto de las puntuaciones obtenidas en la tipificación del año 2000, en la que se constató el efecto Flynn respecto de las normas de 1964. Esta tendencia está en consonancia con investigaciones recientes realizadas en otros países, que muestran un amesetamiento de los puntajes. Un segundo aspecto observado, fue que a medida que aumenta la edad de los sujetos, el rendimiento es mejor. La variabilidad de los puntajes es similar entre los tres grupos de edad analizados.

Palabras clave: efecto Flynn, Matrices Progresivas de Raven, baremos, adolescentes.

Abstract

The current paper presents the results obtained searching for valid and updated norms for Raven's Standard Progressive Matrices, in the city of La Plata, Argentina, and compares them with norms obtained in 1964 and 2000 in order to observe score trends. The sample consisted of 1,001 students of both sexes, aged between 13 and 18 years old, and currently enrolled at one of several high schools. The test was administered collectively, in class-groups, with the examiner's presence, without time limit for its completion, so it would allow the evaluation of intellectual ability without considering the task's speed. Data was split into three age groups (i.e., 13-14 years old, 15-16 years old, and 17-18 years old) and the results were compared with norms from 1964 and 2000. A slight decrease of the direct scores needed to reach an average performance was observed, when compared to those from 2000, which were the norms where the Flynn effect was first noted after 1964. This trend is consistent with recent research from other countries, which shows a plateau effect in direct scores. It was also noticed that as age increases, subjects performance tend to improve. The scores variability is similar in all three age groups.

Keywords: Flynn effect, Raven's Progressive Matrices, norms, adolescents.

Contacto: L. Rossi-Casé. Calle 51 entre 123 y 124, 3er. Piso, Secretaría de Investigaciones. Facultad de Psicología. Ensenada. Provincia de Buenos Aires. Argentina. (C.P. 1925). rossicase@psico.unlp.edu.ar

Cómo citar este artículo:

Rossi-Casé, L., Neer, R., Lopetegui, S., Doná, S. M., Biganzoli, B. y Garzaniti, R. (2014). Matrices Progresivas de Raven: efecto Flynn y actualización de baremos. *Revista de Psicología*, 23(2), 3-13.

<http://dx.doi.org/10.5354/0719-0581.2014.36144>

Introducción

Las pruebas de evaluación psicológica integran el conjunto de instrumentos que permiten al profesional psicólogo arribar a un diagnóstico respecto de aquellos sujetos sobre los que recae la investigación. Estas pruebas aportan información crucial para el desarrollo de estrategias de intervención específicas para cada caso. Las puntuaciones directas que obtienen los sujetos son transformadas a escalas o baremos que permiten su comparación con una muestra estandarizada y dan un sentido estadístico al desempeño individual.

Pero para que la interpretación diagnóstica sea correcta los baremos deben estar actualizados, es decir, el grupo de referencia del cual se obtienen las puntuaciones promedio con las que se compara el rendimiento de un sujeto, debe ser el adecuado. Asimismo, la revisión periódica de los tests en función de los nuevos descubrimientos e hipótesis que aportan las teorías así como los cambios culturales y las adaptaciones regionales, se vuelve necesaria para que el test aplicado sea válido y confiable (Anastasi, 1980).

Un aspecto especial a ser considerado es que los puntajes medios obtenidos en los tests de inteligencia aumentan regularmente y de manera notable en todo el mundo. Este fenómeno se observa desde la misma aparición de los tests. James R. Flynn ha sintetizado los resultados de las investigaciones sobre el tema (Flynn, 1984) y, a raíz de ello, al aumento constante de los valores absolutos en todos los tests de inteligencia se denomina *efecto Flynn*. Lo mencionado hasta aquí advierte que las normas de los tests psicológicos en modo alguno son absolutas, universales o permanentes (Anastasi, 1980).

En este marco, y dada la importancia del Test de Matrices Progresivas de Raven (Raven, J. C., Raven y Court, 2003) como instrumento de evaluación psicológica y su uso generalizado con diferentes propósitos en distintos ámbitos de aplicación de la Psicología, se hace necesaria la actualización de sus baremos para poder utilizarlo de forma confiable. Por ello, el presente trabajo buscó actualizar los baremos del test y comparar los resultados obtenidos con

las normas de 1964 y 2000, para muestras de iguales características con el propósito de verificar la presencia del efecto Flynn en esta población.

Test de Matrices Progresivas de Raven

El Test de Matrices Progresivas de Raven fue publicado por primera vez en el año 1938 por su autor, John C. Raven. Se trata de un test no verbal, de capacidad intelectual, de habilidad mental general. Examina el factor G, para lo cual pone en juego procesos de educación de relaciones y correlaciones sobre un material en el que las variables a considerar no son obvias, es decir, que se deben extraer nuevas comprensiones a partir de la información dada.

Su administración puede ser individual o colectiva, es autoadministrable, sin límite de tiempo en su versión original. Es un test de matrices lacunarias, en el que la tarea del sujeto consiste en completar una matriz, debiendo elegir la respuesta correcta de entre seis u ocho alternativas. Se presenta en varias formas: Escala General, Escala Paralela, Escala Especial Coloreada y Escalas Avanzadas (Raven, J. C. et al., 2003).

La prueba fue revisada en sucesivas oportunidades, generalmente con el fin de obtener normas o baremos adecuados a diferentes poblaciones. En la última versión (Raven, J. C. et al., 2003) se hacen apreciaciones sobre capacidad general, capacidad eductiva, capacidad reproductiva, factor G, inteligencia general, resolución de problemas y competencias, con el sentido de precisar los alcances psicológicos de cada uno de estos términos.

En relación a la capacidad eductiva (Raven, J. C. et al., 2003), esta se define como la capacidad de extraer nuevas comprensiones e información partiendo de lo que se percibe o ya es conocido. Las matrices miden la capacidad para educir relaciones. Esto se debe a que las variables entre las que deben verse relaciones no son en sí mismas obvias y hay que discernir la relación para reconocer esas variables.

La conducta eductiva exige un proceso perceptivo más activo que uno analítico o reproductivo

vo; supone problematizar lo familiar, así como desarrollar la comprensión y resolver problemas. Conceptualmente estaría relacionada con la *inteligencia fluida*, mientras que la capacidad reproductiva se aproxima a la *inteligencia cristalizada* (Cattel, 1968).

La inteligencia fluida se refiere a la habilidad para utilizar el razonamiento abstracto para resolver nuevos problemas que no fueron enseñados en la escuela (Kaufman, 2009). Generalmente se evalúa mediante analogías abstractas cuyos datos no envejecen a lo largo del tiempo ni son específicos de una cultura.

La inteligencia cristalizada, en cambio, es culturalmente más específica, dependiente de la escolarización formal y sujeta a la variedad de los conocimientos adquiridos, por ende, su incremento a lo largo de la vida de una persona es esperable, mientras que el incremento de la inteligencia fluida no lo es (Kaufman, 2009).

Las Matrices Progresivas de Raven son un test de inteligencia fluida, por lo que coincidimos con Fernández Liporace, Ongarato, Saavedra y Casullo (2004) al afirmar que “esto explica la utilización tan difundida de las Matrices en los ámbitos de investigación”, ya que los resultados que aporta permiten comparar poblaciones y/o sujetos que “han sido expuestos a situaciones educativas formales e informales de naturaleza bien disímil” (pp. 50-69).

El efecto Flynn

A partir del uso generalizado de los tests de inteligencia, se ha observado que los puntajes medios obtenidos aumentan regularmente y de manera notable con el correr del tiempo en todo el mundo. James R. Flynn sintetiza los resultados de las investigaciones sobre el tema en el año 1984 y, a raíz de ello, se denomina “efecto Flynn” (Flynn, 1984) al aumento constante de los valores absolutos en todos los tests de inteligencia, para todas las edades.

Esta constatación indica que, para la misma población, las normas para el cociente intelectual se vuelven obsoletas conforme pasan los años. Es decir que, en la actualidad, para obtener el mismo puntaje transformado que hace

algunas décadas, las puntuaciones directas deberían ser más altas, debiendo resolverse un mayor número de problemas.

Estos aumentos son mayores en aquellos tests que miden la inteligencia fluida, que en los que miden la inteligencia cristalizada (Sundet, Barlaug y Torjussen, 2004). Esto llevó a pensar que el aumento está relacionado al factor G de la inteligencia y no a los factores específicos (i.e., factores E).

La causa de este fenómeno permanece desconocida y no existe evidencia suficiente que permita afirmar que este hecho refleje un verdadero aumento de la inteligencia, ligado posiblemente a factores tales como la heterosis, la mejora en las condiciones de vida en las distintas poblaciones, alimentación o nutrición más adecuada; la expansión del sistema educativo; la reducción del tamaño de las familias nucleares, con el consecuente incremento del desarrollo psicológico; la adquisición progresiva de ciertas habilidades para responder satisfactoriamente a los tests; un mayor desarrollo de la escolarización y educación de los niños; la creciente tecnologización de la cultura, desde los juegos de video hasta el acceso cada vez más irrestricto a los medios de comunicación a través de internet, que ofrecen otros aspectos de estimulación (Sundet et al., 2004).

Por su parte, Armstrong y Woodley (2014) afirman que hay evidencia neurológica que señala que el efecto Flynn está asociado a un aumento en el tamaño del cerebro que ha potenciado las funciones del hipocampo, aunque no descartan que otra causa para este fenómeno sea la posibilidad de que los tests que miden la inteligencia fluida y la capacidad educativa, como el Test de Matrices Progresivas de Raven, generen algún tipo de *andamio cognitivo* que permita el uso de habilidades que no dependan del factor G para resolver los problemas.

Debido a la constatación del aumento intergeneracional en los niveles de capacidad educativa, se alcanzó un “efecto de techo” en los adolescentes y adultos jóvenes más capaces, siendo cada vez más frecuente encontrar casos donde el puntaje directo del test era el máximo posi-

ble. Esto introdujo, tras comprobarlo en las estandarizaciones británicas del Test de Raven de 1979, en jóvenes, y de 1992, en adultos, (Raven, J. C. et al., 2003), la necesidad de extender el rango de dificultad de los ítems para devolver al test la capacidad de discriminación que originalmente tenía en estos grupos de sujetos. Así, en la revisión del test de 1998 (Raven, J., Raven y Court, 2003) se incluyeron las formas Paralela y Plus de la Escala General de las Matrices Progresivas.

La Escala General Plus (Raven, J. et al., 2003), que conserva el formato cíclico en 60 ítems de la Escala General Clásica (Raven, J. C. et al., 2003), satisfizo la necesidad de ampliar el rango de dificultad de los ítems sin reducir el poder de discriminación del test entre los sujetos menos capaces (Raven, J. et al., 2003). Asimismo, para los casos en que se necesita una discriminación aún mayor en los extremos inferior y superior de la distribución, se cuenta con la posibilidad de recurrir a las Escalas Coloreada y Avanzada, respectivamente (Raven, J. C. et al., 2003). Simultáneamente, a la Escala General Plus se desarrolló también una versión Paralela (Raven, J. et al., 2003), que al corresponderse con la versión Clásica de 1956 en forma general e ítem por ítem en cuanto a estrategias de solución y dificultad empírica, puede utilizarse con los baremos ya existentes para la Escala General. Por último (Raven, J. et al., 2003), la ubicación de las respuestas correctas en las formas Paralela y Plus de la Escala General fue modificada, para evitar cualquier ventaja que pudiera ofrecer la memorización de las respuestas de la forma Clásica.

En su estudio inicial, Flynn (1984) estableció que la magnitud del aumento del cociente intelectual –medido inicialmente con las escalas Wechsler y Stanford-Binet– era de 0.3 puntos por año o 3 puntos por década. Estos estudios se realizaron comparando muestras estadounidenses entre los años 1932 y 1978. Más adelante, la misma tendencia se corroboró con datos de otros 20 países.

No obstante, investigaciones recientes muestran una disminución en el crecimiento de dichos puntajes. Este efecto fue observado por primera vez en los países escandinavos. Tal es

el caso de investigaciones longitudinales hechas en Noruega (Sundet et al., 2004) en las cuales se muestra una disminución en los puntajes de una batería de tests administrados a los jóvenes de 18 años, antes de entrar al servicio militar, desde los años 50. Las pruebas utilizadas son tests de matemática y lenguaje, similares a los subtests del WAIS IV, y un test no verbal que fue construido de forma similar al test de Raven. Los dos primeros miden la inteligencia cristalizada, mientras que el último mide la inteligencia fluida.

Teasdale y Owen (2007) estudiaron los datos aportados por una batería similar en Dinamarca, la cual desde 1957 se administra a todos los jóvenes de 18 años que ingresan al servicio militar. De los cuatro tests que la componen, los autores afirman que el de Matrices de Letras se asemeja a las Matrices Progresivas de Raven. Los autores encontraron que si bien hubo un pequeño incremento de los puntajes entre 1988 y 1998, los mismos disminuyeron en 2003-2004, incluso por debajo de los obtenidos en 1988 para todos los tests de la batería. Esto se observó en jóvenes de todos los niveles educativos. Con esto, los autores apoyan la afirmación de que en lo que va del siglo XXI ha habido muy poca evidencia sobre la continuidad del efecto Flynn, tal como se lo define hasta este momento, es decir, como el aumento continuo de las medias en los tests de inteligencia.

Por su parte, Brouwers, Van de Vijver y Van Hemert (2008) hicieron una revisión de 193 estudios sobre las tres escalas Raven publicados entre 1944 y 2003, que abarcan 798 muestras provenientes de 45 países, siendo el número total de sujetos 244316. Dentro de ese reporte, Argentina se encuentra representada por más de 20 estudios publicados en dicho período. Los autores calcularon la correlación entre las características de las muestras (edad, años de escolaridad), de los países de los que provenía cada estudio (Producto Bruto Interno [PBI], escolaridad y analfabetismo), y de los estudios consultados (año de publicación). Encontraron que, si bien tanto el PBI como los años de escolaridad se correlacionan positivamente con el desempeño en el test (por lo que pueden estar influyendo en las diferencias interculturales en el desempeño), existe una correlación

negativa significativa entre el PBI y la magnitud del efecto Flynn. A partir de esto, concluyen que el efecto Flynn habría alcanzado su techo en los países desarrollados de occidente, mientras que los países con un PBI per cápita menor, todavía muestran un efecto Flynn más pronunciado.

Esta conclusión coincide con la de Teasdale y Owen (2007), quienes afirman que si el efecto Flynn llega a su fin en los países más desarrollados, estaría lejos de hacerlo en el resto del mundo, pero que se esperaría que estas diferencias entre los distintos países disminuyeran gradualmente en el futuro.

Flynn (2013) ha resaltado que también Holanda y Finlandia muestran una disminución en la magnitud del efecto que lleva su nombre, sin embargo, esta tendencia no se evidencia en todos los países desarrollados. Los datos recientes provenientes de Australia son ambiguos al respecto y Estados Unidos y Corea del Sur constituyen claras excepciones a la misma. También menciona que algunos autores atribuyeron esta tendencia a la inmigración de poblaciones con CI bajos en los países desarrollados, pero que esto no explicaría la situación de Estados Unidos. Flynn aventura la idea de que la detención de los incrementos del CI en ciertos países podría deberse a que en ellos ya no puede progresarse más en aquellos aspectos que se supone fomentarían dichos incrementos.

Rusell (2007) sostiene que un efecto meseta sería de esperar en culturas que tengan condiciones de vida óptimas, lo cual explicaría que dicho efecto se haya manifestado en los países escandinavos antes que en el resto del mundo, ya que estos han establecido un sistema de bienestar que incluye a toda la población desde el fin de la Segunda Guerra Mundial. Recurriendo a las escalas Wechsler y empleando la regresión lineal, este autor estima que en Estados Unidos el efecto meseta se alcanzará en 2024. Pero al considerar posibles sesgos en la selección de las muestras para las estandarizaciones previas de estas escalas, estima que la meseta ya podría haberse alcanzado en 2004, por lo que el efecto Flynn podría ya no estar vigente en ese país.

Con respecto a la Argentina, este equipo realizó una primera constatación del efecto Flynn en la ciudad de La Plata y alrededores al actualizar los baremos del Test de Matrices Progresivas de Raven en el año 2000, y compararlos con la estandarización previa, del año 1964. Esta comparación mostró un considerable aumento en los puntajes directos del test para todas las edades (Flynn y Rossi-Casé, 2012; Rossi-Casé, Neer y Lopetegui, 2001, 2002, 2011). Asimismo, se realizó un análisis de los resultados por sexo y los resultados sirvieron para comprobar tendencias a nivel internacional (Flynn y Rossi-Casé, 2011).

Método

Participantes

La muestra de estandarización consistió en 1001 sujetos de ambos sexos, con edades que variaban de los 13 a los 18 años, considerando la edad en años cumplidos al momento de la administración del test. En el presente estudio se dividió el rango de edad de 13 a 18 años, en tres intervalos iguales, teniendo en cuenta que el desempeño en la prueba puede cambiar bastante en cada uno de esos intervalos, lo cual justifica la generación de normas diferentes para cada grupo de edades.

La muestra fue estratificada con respecto a edad, sexo y tipo de gestión de los colegios (pública o privada), utilizando los datos del censo poblacional del año 2010 (Instituto Nacional de Estadística y Censos [Indec], 2010). Se utilizó un muestreo bietápico que permitió seleccionar distintos conglomerados (instituciones educativas) y en cada uno de ellos seleccionar aleatoriamente los grupos de clase que serían examinados para las distintas edades, siendo evaluados todos los sujetos que conformaban cada grupo.

Trabajando con un intervalo de confianza del 95% se pudo establecer que el error muestral máximo sería del 3% para este tamaño de muestra. Las instituciones educativas fueron tanto públicas como privadas, todas pertenecientes a la educación secundaria y ubicadas en la ciudad de La Plata. A continuación, en

la tabla 1, se describe la composición de la muestra, según sexo y edad.

Instrumentos

Se utilizó el Test de Matrices Progresivas de Raven; Escala General, segunda edición, 2003. El test está compuesto por 60 problemas organizados en cinco series (A; B; C; D; E), de doce ítems cada uno. Cada una de las series comienza con problemas fáciles y va aumentando la complejidad a lo largo de la misma.

Procedimientos

El test se administró de manera colectiva, en grupos-clase, con presencia del examinador y sin límite de tiempo para su ejecución, para así permitir la evaluación de la capacidad intelectual sin la intervención de la velocidad en la tarea. Para la realización de esta tarea se solicitó, en tiempo y forma, la debida autorización a los Inspectores Educativos del Distrito La Plata y se obtuvo la autorización previa de los padres o tutores de los adolescentes para su participación en la investigación mediante un formulario de consentimiento informado, donde se explicaba el propósito de la investigación y la confidencialidad de los datos.

El test fue administrado en distintos períodos de los años 2012 y 2013 (noviembre de 2012, junio, octubre y noviembre de 2013).

La administración estuvo a cargo de profesionales y alumnos avanzados de la carrera de Psicología, entrenados previamente a fin de homogeneizar el procedimiento. Para la consigna, se siguieron de manera rigurosa las instrucciones dadas por el autor del test para la administración colectiva. Se administró en cada una de las escuelas seleccionadas, en horarios regulares de clases, cedidos por las autoridades y profesores de cada institución. Cada grupo no superó los 30 alumnos.

Al conformar la base de datos, se excluyeron aquellos protocolos que mostraban una discrepancia mayor a la esperada en la composición de los puntajes ($n = 144$). Se analizó una base de datos que incluye 1001 protocolos.

Análisis

Con la información obtenida se conformó una base en el programa Excel (versión 12.0.4518.1014) obteniéndose los estadísticos descriptivos correspondientes así como los percentiles que nos permitieron elaborar las normas para este grupo.

Los resultados obtenidos en el estudio actual se compararon con las normas obtenidas en los años 1964 y 2000, para los mismos grupos etarios de la ciudad de La Plata, Argentina, utilizando el estadístico t de Student con un nivel de confianza de 95%.

Tabla 1
Composición de la muestra

Edad	Sexo		Total de casos
	Mujer	Hombre	
Totales por grupo de 13-14 años	141	220	361
13 años	73	135	208
14 años	68	85	153
Totales por grupo de 15-16 años	165	171	336
15 años	82	99	181
16 años	83	72	155
Totales por grupo de 17-18 años	130	174	304
17 años	84	94	178
18 años	46	80	126
Total de casos por sexo	436	565	1001

El análisis de datos se realizó por grupos de edad: 13-14 años; 15-16 años y 16-17 años, para poder comparar sus resultados con otros estudios de la misma índole (ver resultados en apéndice)

Resultados

La descripción de las puntuaciones directas obtenidas muestra que en los tres grupos de edad se alcanzó la puntuación máxima posible. Las puntuaciones mínimas resultaron similares para todas las edades. Los resultados, expuestos en la tabla 2, mostraron que a medida que aumenta la edad de los sujetos, el rendimiento de los grupos mejora y la heterogeneidad de sus respuestas disminuye.

Tabla 2

Descripción de las puntuaciones directas, por rango de edades

Grupo de edades	M	DE	Rango de respuestas
13-14	44.01	7.60	12 a 60
15-16	48.39	6.75	10 a 60
17-18	49.15	6.57	12 a 60

Nota: $n = 1001$. Puntuación máxima: 60 puntos.

Resultados para el grupo de edad 13-14 años

En los sujetos que integran este grupo se observan los resultados más bajos de todos los que conforman la muestra. La puntuación promedio es de 44.05 puntos y la desviación estándar es de 7.60 puntos. El rango de respuestas correctas fue de 12 para los más bajos, y de 60 para los más altos.

La mitad de los sujetos que integran este grupo de edad obtuvo 45 puntos o menos. Si bien esto representa un incremento de 6 puntos respecto del baremo del año 1964, el resultado observado es 3 puntos menor que la norma alcanzada en el año 2000. Este comportamiento de un incremento significativo de las puntuaciones respecto de las normas 1964, $t(359) = 1.98$; $p < .05$ y una disminución respecto de los baremos año 2000, se observó en todos los valores percentilares calculados. Los resultados se muestran en la tabla 3.

Tabla 3

Comparación de las puntuaciones directas, años 1964, 2000 y 2013 para el rango de edad entre 13-14 años

Percentil	Baremos		Valores percentilares
	1964	2000	2013
P99	50	59	57
P95	49	56	54
P90	48	55	52
P75	44	52	50
P50	39	48	45
P25	35	44	40
P10	27	39	34
P5	22	36	30

Nota: Elaboración propia basada en la administración del Test de Raven-Escala General (Raven, J. C. et al., 2003), en 2012-2013, a estudiantes secundarios de ambos sexos, en La Plata, Argentina ($n = 361$).

Resultados para el grupo de edad 15-16 años

La puntuación promedio observada es de 48.39 puntos y la desviación estándar es de 6.75 puntos. El rango de respuestas correctas fue de 10 para los más bajos, y de 60 para los más altos.

La mitad de los sujetos que integran este grupo obtuvo al menos 50 puntos. Esto representa un incremento de 9 puntos respecto del baremo del año 1964, y 1 punto respecto de la norma alcanzada en el año 2000. Este incremento significativo de las puntuaciones respecto de las normas 1964 se observó en todos los valores percentilares calculados, $t(334) = 1.41$; $p < .05$. La comparación de resultados con los baremos del año 2000 muestra aumento de 1 y 2 puntos en los percentiles 25, 50 y 75, mantenimiento de los puntajes para los percentiles 10, 90 y 95; y una disminución de 1 punto en percentiles 5 y 99. Los resultados se muestran en la tabla 4.

Resultados para el grupo de edad 17-18 años

En los sujetos que integran este grupo se observan los resultados más altos de todos los que

integran la muestra. La puntuación promedio es de 49.15 puntos y la desviación estándar de 6.57 puntos. El rango de respuestas correctas fue de 12 para los más bajos y de 60 para los más altos.

Tabla 4

Comparación de las puntuaciones directas, años 1964, 2000 y 2013 para el rango de edad entre 15-16 años

Percentil	Baremos		Valores percentilares
	1964	2000	2013
P99	52	59	58
P95	51	56	56
P90	49	55	55
P75	44	52	53
P50	41	49	50
P25	36	44	46
P10	28	40	40
P5	23	36	35

Nota: Elaboración propia basada en la administración del Test de Raven-Escala General (Raven, J. C. et al., 2003), en 2012-2013, a estudiantes secundarios de ambos sexos, en La Plata, Argentina ($n = 336$).

La mitad de los sujetos obtuvo 50 puntos o menos. El resultado observado representa un incremento de 10 puntos respecto del baremo del año 1964, al tiempo que permanece igual a la norma alcanzada en el año 2000. Este incremento significativo de las puntuaciones respecto de las normas 1964, $t(302) = 0.64$; $p < .05$ se observó en todos los valores percentilares calculados. Respecto de los baremos del año 2000, se constata la disminución de 3 y 2 puntos para los valores percentiles 5 y 10; igualdad en las puntuaciones de los percentiles 25, 50, 90 y 95; y un aumento de 1 punto en los valores de los percentiles 75 y 99. Los resultados se muestran en la tabla 5.

Resultados para el grupo de edad 13-16 años y 17-18 años

La metodología aplicada para la descripción de los resultados obtenidos en el baremo del año 2000 permite mostrar que el rendimiento promedio de los sujetos de 13 a 16 años en el estudio actual es 1.64 puntos menor que aquel; y

la variabilidad de las respuestas es 1.33 puntos mayor en la actualidad. Para los sujetos del grupo de edad 17-18 años, los rendimientos promedio son levemente inferiores, y la variabilidad aumenta en 1.13 puntos. Los resultados se muestran en la tabla 6.

Tabla 5

Comparación de las puntuaciones directas, años 1964, 2000 y 2013 para el rango de edad entre 17-18 años

Percentil	Baremos		Valores percentilares
	1964	2000	2013
P99	53	58	59
P95	53	57	57
P90	50	56	56
P75	45	53	54
P50	40	50	50
P25	35	46	46
P10	27	43	40
P5	22	40	38

Nota: Elaboración propia basada en la administración del Test de Raven-Escala General (Raven, J. C. et al., 2003), en 2012-2013, a estudiantes secundarios de ambos sexos, en La Plata, Argentina ($n = 304$).

Tabla 6

Comparación de las puntuaciones directas, años 2000 y 2013, por grupos de edades

Grupos de edad	Puntuaciones directas			
	2000		2013	
	M	DE	M	DE
13-16	47.76	6.19	46.12	7.52
17-18	49.50	5.44	49.15	6.57

Nota: Elaboración propia basada en la administración del Test de Raven-Escala General (Raven, J. C. et al., 2003), en 2012-2013, a estudiantes secundarios de ambos sexos, en La Plata, Argentina ($n = 1001$).

Discusión y conclusiones

Respecto de las normas del año 1964, los resultados observados en esta muestra corroboran el efecto Flynn ya constatado con los baremos del año 2000. Sin embargo, entre el baremo 2000 y el estudio actual, no se observa dicho efecto.

En consonancia con investigaciones recientes realizadas en otros países (Brouwers, Van de Vijver y Van Hemert, 2008; Rusell, 2007; Sundet et al., 2004; Teasdale y Owen, 2007), los resultados obtenidos permiten mostrar una detención del aumento de los puntajes directos necesarios para alcanzar el rendimiento promedio en cada rango de edad. En el estudio actual se constata una leve disminución de las puntuaciones promedio que no resulta significativa, por lo que hablamos de un amesetamiento de los puntajes.

Si el efecto Flynn establece que los puntajes han ido creciendo a razón de más o menos una desviación estándar por generación (Flynn, 1984), resulta adecuado preguntarse si este aumento sería indefinido en el tiempo y, en ese caso, cuándo cesaría. Para intentar dar respuesta a esta cuestión deberíamos volver a analizar las hipótesis que en su momento intentaron explicar la presencia del efecto Flynn, tales como mejoras en la alimentación, mayor escolarización, el uso de las nuevas tecnologías y el predominio de lo visual.

Observamos que la presencia de estos estímulos se volvió común y generalizada en este lapso, y que los participantes de este estudio, adolescentes con edades entre 13 y 18 años, son en su mayoría nativos tecnológicos de la aldea global. Al mismo tiempo, la facilidad de acceso a la información generaliza la difusión

de los conocimientos y avances relacionados con la salud y la nutrición (Rossi-Casé, Neer y Lopetegui, 2001).

Entonces, si las condiciones que alguna vez se mencionaron para explicar el efecto Flynn se han estabilizado en los últimos años, los puntajes, que eran su expresión, han corrido la misma suerte. No habiendo ninguna otra condición que produzca un cambio cualitativo del mismo tenor, las puntuaciones directas se mantienen estables, lo que condice con el amesetamiento observado en Argentina y otros países del mundo.

Otra posible explicación de este fenómeno podría ser que el test de Matrices Progresivas de Raven no estuviese midiendo únicamente lo que dice medir, y que hubiera otros factores que incidan en la resolución del test, como por ejemplo, la educación formal y, por ende, la inteligencia cristalizada. Estos factores podrían ser más sensibles a los estímulos del ambiente, lo que podría explicar no solo los grandes cambios que se dan a través de los años, sino también las diferencias que existen entre los países desarrollados y aquellos en vías de desarrollo.

Estas posibles explicaciones son solo intentos por profundizar la comprensión de las causas del amesetamiento.

Referencias

- Anastasi, A. (1980). *Tests psicológicos*. Madrid: Aguilar.
- Armstrong, E. L. y Woodley, M. A. (2014). The rule-dependence model explains the commonalities between the Flynn effect and IQ gains via retesting. *Learning and Individual Differences*, 29, 41-49.
<http://dx.doi.org/10.1016/j.lindif.2013.10.009>
- Brouwers, S. A., Van de Vijver, F. J. R. y Van Hemert, D. A. (2008). Variation in Raven's Progressive Matrices scores across time and place. *Learning and Individual Differences*, 19, 330-338.
<http://dx.doi.org/10.1016/j.lindif.2008.10.006>
- Cattel, R. B. (1968). The theory of fluid and crystallized intelligence. Its relationship to culture free tests and its verification in 9-12 ys. old children. *Bolletino di Psicologia Applicato*, 88(90), 3-22.
- Fernández Liporace, M., Ongarato, P., Saavedra, E. y Casullo, M. M. (2004). El Test de Matrices Progresivas, Escala General: un análisis psicométrico. *Evaluar*, 4, 50-69.
- Flynn, J. R. (1984). The mean IQ of Americans: massive gains 1932-1978. *Psychological Bulletin*, 95, 29-51.
- Flynn, J. R. (2013). The "Flynn Effect" and

- Flynn's paradox. *Intelligence*, 41(6), 851-857. <http://dx.doi.org/10.1016/j.intell.2013.06.014>
- Flynn, J. R. y Rossi-Casé, L. (2011). Modern women match men on Raven's Progressive Matrices. *Personality and Individual Differences*, 50(6), 799-803. <http://dx.doi.org/10.1016/j.paid.2010.12.035>
- Flynn, J. R. y Rossi-Casé, L. (2012). IQ gains in Argentina between 1964 and 1998. *Intelligence*, 40, 145-150. <http://dx.doi.org/10.1016/j.intell.2012.01.006>
- Instituto Nacional de Estadística y Censos. (2010). *Censo Nacional de Población, Hogares y Viviendas 2010*. Recuperado de <http://www.censo2010.indec.gov.ar/>
- Kaufman, A. S. (2009). Hot topic: Are our IQs fixed or are they malleable? En J. C. Kaufman (Ed.), *IQ Testing 101* (pp. 201-222). New York: Springer.
- Raven, J., Raven, J. C. y Court, J. H. (2003). *Standard Progressive Matrices. Manual, Section 3*. Oxford: Oxford Psychologists Press.
- Raven, J. C., Raven, J. y Court, J. H. (2003). *Test de Matrices Progresivas: Escala General*. Buenos Aires: Paidós.
- Rossi-Casé, L., Neer, R. y Lopetegui, S. (2001). Test de Matrices Progresivas de Raven: comparación de baremos. El aumento de los puntajes directos a través del tiempo. *Evaluar*, 2(2), 39-51.
- Rossi-Casé, L., Neer, R. y Lopetegui, S. (2002). Test de Matrices Progresivas de Raven: construcción de baremos y constatación del "Efecto Flynn". *Orientación y Sociedad*, 3, 181-187.
- Rossi-Casé, L., Neer, R. y Lopetegui, S. (2011). Baremo ciudad de La Plata (2000). Buenos Aires, Argentina. Escala General (MPG). En J. C. Raven (Ed.), *Test de Matrices Progresivas. Carpeta de evaluación, Escala General* (pp. 19-23). Buenos Aires: Paidós.
- Rusell, E. W. (2007). Commentary: The Flynn effect revisited. *Applied Neuropsychology*, 14(4), 262-266. <http://dx.doi.org/10.1080/09084280701719211>
- Sundet, J. M., Barlaug, D. G. y Torjussen, T. M. (2004). The end of the Flynn effect? A study of secular trends in mean intelligence test scores of Norwegian conscripts during half a century. *Intelligence*, 32, 349-362. <http://dx.doi.org/10.1016/j.intell.2004.06.004>
- Teasdale, T. W. y Owen, D. R. (2007). Secular declines in cognitive test scores: A reversal of the Flynn effect. *Intelligence*, 36, 121-126. <http://dx.doi.org/10.1016/j.intell.2007.01.007>

Fecha de recepción: 14 de marzo de 2014
 Fecha de aceptación: 29 de abril de 2014

Apéndice

Normas 2014. Test de Raven. Escala General. La Plata, Argentina.

Percentil	Puntuación directa según edades		
	13-14 años	15-16 años	17-18 años
P99	57	58	59
P95	54	56	57
P90	52	55	56
P75	50	53	54
P50	45	50	50
P25	40	46	46
P10	34	40	40
P5	30	35	38