



Asociación de Psicología de Puerto Rico

PO Box 363435 San Juan, Puerto Rico 00936-3435

Tel. 787.751.7100 Fax 787.758.6467

www.asppr.net E-mail: info@asppr.net

Revista Puertorriqueña de Psicología
Volumen 20, 2009

Validación concurrente de la *Wechsler Intelligence Scale for Children-Fourth Edition, Spanish* con la Escala de Inteligencia Wechsler para Niños-Revisada de Puerto Rico^{1 2}

**Patricia Torres Díaz, Ivonne Romero García,
Walter Rodríguez Irizarry y Carmen Rodríguez Fernández**
*Universidad Interamericana de Puerto Rico,
Recinto de San Germán*

Resumen

Examinamos la validez concurrente de la *Wechsler Intelligence Scale for Children-Fourth Edition, Spanish (WISC-IV Spanish)* con la Escala de Inteligencia Wechsler para Niños-Revisada de Puerto Rico (EIWN-R PR) con una muestra de 40 estudiantes entre las edades de 11 y 13 años. Se encontraron correlaciones positivas y significativas entre las puntuaciones a escala y las puntuaciones estándares de ambos instrumentos. Los coeficientes de correlación fluctuaron entre .45 y .81. Se obtuvo una correlación estadísticamente significativa de $r = .63$ entre las puntuaciones de la Escala Total. Al comparar los resultados de la Escala Total en ambos instrumentos, se encontraron puntuaciones 14 puntos por debajo en la *WISC-IV Spanish*. Esta diferencia puede ser explicada por factores asociados al error estándar de la medida, cambios normativos y estructurales entre los dos instrumentos, tamaño de la muestra y el efecto Flynn. Estos factores y las interacciones entre ellos deben tomarse en consideración al interpretar los resultados de

¹ *Nota:* Este artículo fue sometido a evaluación en septiembre de 2008 y aceptado para publicación en abril de 2009.

² Los autores desean agradecer la labor editorial del Dr. Juan R. González Mendoza en la redacción de este artículo. También agradecen a aquellos niños que generosamente aceptaron participar de esta investigación. La correspondencia sobre este artículo debe enviarse a Patricia Torres, PhD, Departamento de Psicología, Universidad Interamericana de Puerto Rico, Recinto de San Germán, PO Box 5100, San Germán, Puerto Rico 00683. E-mail: ptorres92@yahoo.com

este estudio. Recomendamos investigaciones con una muestra más amplia y representativa de niños puertorriqueños.³

Palabras claves: *WISC-IV Spanish, EIWN-R PR, validez concurrente, medición de CI*

Abstract

We examined the concurrent validity of the Wechsler Intelligence Scale for Children-Fourth Edition, Spanish (WISC- IV Spanish) using the *Escala de Inteligencia Wechsler para Niños – Revisada de Puerto Rico (EIWN-R PR)* as a criterion measure. The subjects were 40 students between the ages of 11 and 13. Statistically significant and positive correlation coefficients were found between the scale and standard scores of both instruments; ranging between .45 and .81. A significant correlation coefficient of $r=.63$ was found between the Full Scale IQ's. Comparisons between the Full Scale IQ's of both instruments revealed *WISC-IV Spanish* total scores that were 14-points lower. This difference may be explained by factors associated with the standard error of measurement, normative and structural changes between the two instruments, sample size, and the Flynn effect. These factors and the interactions with each other should be considered when interpreting the results of this study. We suggest further research with a larger and more representative sample of Puerto Rican children.

Keywords: *WISC-IV Spanish, EIWN-R PR, concurrent validity, measurement of IQ*

La *Wechsler Intelligence Scale for Children-Fourth Edition, Spanish* (a partir de ahora la *WISC-IV Spanish*) se publicó en 2005, para utilizarse con la población de “niños hispano-parlantes entre las edades de 6 años 0 meses hasta los 16 años 11 meses que están aprendiendo inglés como segundo idioma, muchos de los cuales están en el proceso de aculturación a Estados Unidos” (Wechsler, 2005, p.1; cita traducida al español por los autores de este artículo).

Aunque en el Manual del instrumento se indica que la *WISC-IV Spanish* fue diseñada primordialmente para los niños latinos en

³ En este documento se reconoce la importancia de atender apropiadamente las expresiones relacionadas al género de las personas, no obstante, para facilitar su lectura, se utilizará la expresión masculina para referirnos a las personas de ambos géneros.

Estados Unidos, en Puerto Rico ya ha comenzado su utilización. Esta práctica genera múltiples controversias e interrogantes sobre la equivalencia de las puntuaciones obtenidas en la *WISC-IV Spanish* y la Escala de Inteligencia Wechsler para Niños-Revisada de Puerto Rico (a partir de ahora la *EIWN-R PR*, Wechsler, 1992). Esta última precedió a la *WISC-IV Spanish* y fue normalizada con una muestra representativa de niños y adolescentes puertorriqueños a finales de los años 80 y principios de los 90 (Wechsler, 1992, p.8).

Actualmente, se ha cuestionado la validez y confiabilidad de la *WISC-IV Spanish* y su aplicabilidad en la evaluación cognoscitiva de niños puertorriqueños. También, han surgido varios cuestionamientos, de naturaleza ética y técnica, en torno a la equivalencia de las puntuaciones obtenidas con el nuevo instrumento, la pertinencia de utilizar una nueva escala que incorpora adelantos en la medición del procesamiento de información, pero que carece de normas para niños puertorriqueños, y las implicaciones del uso de una prueba como la *EIWN-R PR*, con datos normativos de hace aproximadamente 20 años. Estos asuntos serán discutidos por separado más adelante en el texto de este artículo.

La *WISC-IV Spanish*, derivada de la *Wechsler Intelligence for Children-Fourth Edition (WISC-IV; Wechsler, 2003)*, es una prueba que representa un cambio sustancial de enfoque y estructura sobre sus predecesoras. Incorpora adelantos en la evaluación neuropsicológica, como por ejemplo, la medición de funciones ejecutivas (Wechsler, 2005; Flanagan & Kaufman, 2004). Además, introduce nuevos índices y subpruebas para mejorar la medida de razonamiento fluido, el procesamiento de información, el almacenamiento temporal de la información y la manipulación de tareas cognitivas más complejas (Burns & O'Leary, 2004). También, añade y cuantifica los principios de la evaluación de procesos propuestos por Kaplan (1988) en algunas de las subpruebas, lo que permite al examinador un análisis cualitativo de la ejecución del niño.

La *WISC-IV Spanish*, incorpora el enfoque de procesamiento de información que permite cumplir con la ley federal IDEIA (PL-108-466) del 4 de diciembre de 2004, en su definición de problemas específicos de aprendizaje. Esta Ley requiere que se evalúen los *procesos psicológicos básicos* y se utilicen medidas precisas, más allá de ofrecer puntuaciones globales de coeficiente intelectual (CI), que permitan evaluar de forma exhaustiva, los múltiples factores cognoscitivos invo-

lucrados en la llamada *conducta inteligente*. Por tanto, el psicólogo puertorriqueño se encuentra ante la disyuntiva ética de utilizar un instrumento en español específicamente desarrollado para evaluar procesos psicológicos múltiples, que no incluye normas para niños en Puerto Rico.

El otro aspecto de la controversia propone que las puntuaciones obtenidas con la *EIWN-R PR* podrían sobreestimar el funcionamiento intelectual de los niños evaluados, porque los datos normativos provienen de una muestra constituida por estudiantes puertorriqueños matriculados en escuelas públicas y privadas de Puerto Rico “durante los años 1986-87 y 1989-90” (Wechsler, 1992, p.8). Se plantea que las transformaciones culturales, sociales y tecnológicas en Puerto Rico y otros países, producto del paso de los años, resultan en cambios que podrían alterar las puntuaciones en este tipo de prueba (Morris & Maisto, 2001; Pons, 2007).

Estos planteamientos están documentados ampliamente en la literatura científica y son denominados como el efecto Flynn. Este efecto establece que en las pruebas de inteligencia cristalizada como las Escalas Wechsler, se exhiben cambios en el CI de aproximadamente nueve puntos por generación, o sea, cambios de nueve puntos en un lapso de 30 años (Flynn, 1994). Según Flynn (1987), en las pruebas que miden inteligencia fluida, como las pruebas de inteligencia no verbal, se pueden encontrar aun mayores cambios, de cinco a 25 puntos en el CI por generación. Es importante señalar, que Flynn y sus colegas expresan los índices de cambio en .308 puntos de CI por año, así que multiplicados por 30 años, proveen un aumento estimado durante una generación (Flynn, 1987, 1994, 1999; Flynn & Weiss, 2007).

El efecto Flynn pone de relieve la necesidad de que las normas de las pruebas psicológicas se reevalúen y revisen cada 15 a 20 años, ya que éstas se vuelven obsoletas y hay que ajustar la puntuación promedio a 100 (Flynn, *et al.*, 2007). En una investigación realizada por Kanaya, Scullin y Ceci (2003) se compararon las puntuaciones de la *Wechsler Intelligence Scale for Children-Revised* (WISC-R; Wechsler, 1974) con la *Wechsler Intelligence Scale for Children-Third Edition* (WISC-III; Wechsler, 1991). Los resultados revelaron que el CI obtenido en la versión más reciente del instrumento fue más bajo, variando en promedio 5.6 puntos. Estos autores alertan sobre el impacto del efecto Flynn en diferentes contextos, tales como el de educación especial y el sistema legal, donde se toman decisiones sobre la elegibilidad

de servicios a partir de las normas de las pruebas psicológicas.

Aunque no se han hallado investigaciones que comparen las puntuaciones obtenidas en la *WISC-IV Spanish* y la *EIWN-R PR*, se revisaron varios estudios realizados en Estados Unidos con ediciones anteriores y al inglés de estos instrumentos. Los estudios realizados con la *WISC-R* y *WISC-III* con niños excepcionales (talentosos y condiciones especiales, respectivamente) reportaron correlaciones positivas y significativas, las cuales fluctuaron entre .76 y .74 en la Escala Verbal, entre .56 y .66 en la Escala de Ejecución y entre .85 y .66 en la Escala Total (Lyon, 1995; Sevier & Bain 1994). Las investigaciones con la *WISC-III* y la *WISC-R* también revelaron discrepancias de ocho a 14 puntos entre la Escala Total de ambas pruebas (Carlton & Sapp, 1997; Jones, 1996; Sevier & Bain, 1994).

En vista de la controversia sobre la validez y confiabilidad de las evaluaciones que utilizan la *WISC-IV Spanish* en Puerto Rico, es importante explorar si los resultados obtenidos con esta prueba son comparables con los de instrumentos adaptados y normalizados para la población bajo consideración. Por tanto, esta investigación evaluó las puntuaciones obtenidas en la *WISC-IV Spanish* para compararlas con las puntuaciones obtenidas en la *EIWN-R PR*, instrumento considerado como medida criterio. Para ello se utilizó una muestra de niños puertorriqueños que habitan la región sur de Puerto Rico.

Método

Participantes

Participaron en esta investigación 40 estudiantes, 21 (52%) del género femenino y 19 (48%) del género masculino, entre las edades de 11 y 13 años, alumnos de una escuela privada de la región de Ponce. Éstos cursaban el sexto (18%), séptimo (53%) y el octavo grado (40%). Los participantes fueron seleccionados por disponibilidad y participaron voluntariamente. Se utilizaron los criterios de exclusión, señalados para la muestra de estandarización del *WISC-IV Spanish*. Se excluyeron los estudiantes referidos o evaluados a educación especial y los que tenían historial de problemas físicos o mentales.

Instrumentos

Para esta investigación se utilizaron los siguientes instrumentos:

la *WISC-IV Spanish* y la *EIWN-R PR*. También, se administró a los padres o encargados legales una hoja de datos confidenciales sobre el nivel de educación alcanzado por ellos como indicador del nivel socioeconómico.

La *WISC-IV Spanish* elimina la Escala Verbal (*EV*) y la de Ejecución (*EE*), característica de los instrumentos predecesores. Dicho instrumento utiliza cinco puntuaciones compuestas: el Cociente Intelectual Total (*FSIQ*) y cuatro índices de funcionamiento. Éstos son: *Índice de Comprensión Verbal (VCI)*, *Índice de Razonamiento Perceptual (PRI)*, *Índice de Memoria de Trabajo (WMI)* e *Índice de Velocidad de Procesamiento (PSI)* (Wechsler, 2005). El *Índice de Comprensión Verbal* provee un estimado de las habilidades verbales que usan razonamiento, comprensión y formación de conceptos. El *Índice de Razonamiento Perceptual* evalúa el razonamiento perceptual y la organización. El *Índice de Memoria de Trabajo* es uno nuevo que evalúa primordialmente las funciones ejecutivas. La tarea requiere almacenar información en memoria a corto plazo, manipular mentalmente la información registrada y emitir una respuesta (Baddeley, 2003; Matlin, 2005; Prifitera, Saklofske, & Weiss, 2005). Además de requerir un alto grado de concentración y atención, el nuevo *Índice de Velocidad de Procesamiento* evalúa la rapidez en que se procesa la información (Wechsler, 2005; Weiss, Saklofske, Prifitera, & Holdnack, 2006).

La estandarización de la *WISC-IV Spanish* se realizó con 851 niños de padres latinos que provenían de: México, Cuba, Puerto Rico, Centro y Sur América y la República Dominicana (Wechsler, 2005). Cuatro regiones conformaron la muestra: el noroeste, sur, noreste y este de Estados Unidos. También, se incluyó a niños residentes en Puerto Rico en la muestra. Éstos fueron clasificados como pertenecientes a la región sur de Estados Unidos. Según los datos provistos por María Muñoz, Analista de Investigación del *Psychological Corporation* (comunicación vía correo electrónico, 10 de mayo de 2007), de ese grupo de niños latinos, 55 eran de origen puertorriqueño. Según Muñoz, aunque la muestra era pequeña, constituía cerca del 10% de la muestra de confiabilidad de la *WISC-IV Spanish*. Sin embargo, del grupo de niños de origen puertorriqueño, sólo 18 residían en Puerto Rico.

La *EIWN-R PR* se escogió como medida de criterio por ser una prueba válida, confiable, estandarizada y normalizada con una muestra

representativa de 2,200 niños puertorriqueños (Wechsler, 1992). Los índices de confiabilidad de esta prueba sobrepasan .91 para todas las subpruebas y Escalas. Su validez ha sido documentada ampliamente en varias investigaciones (Herrans & Rodríguez, 1992).

La administración de la *EIWN-R PR* permite obtener tres puntuaciones: *Cociente Intelectual Total (CIT)*, *Cociente Intelectual Verbal (CIV)* y un *Coficiente Intelectual de Ejecución (CIE)*. El análisis de factores mostró consistencia y estabilidad en los factores de *Comprensión Verbal*, *Organización Perceptual* y *Factor de Libertad de Distracción* (Carroll, Herrans & Rodríguez, 1995). El factor de *Comprensión Verbal* evalúa el conocimiento verbal obtenido por medio de la educación formal e informal. El factor de *Organización Perceptual* provee un estimado de la habilidad de interpretar y organizar material percibido de manera visual dentro de un límite de tiempo. El factor de *Libertad de Distracción* evalúa la capacidad del niño para modular su atención, concentración, memoria y ejercer control mental (Kaufman, 1975, 1982). En la Tabla 1 presentamos una descripción de la estructura de los dos instrumentos. Identificamos, además, los factores de la *EIWN-R PR*.

TABLA 1
Descripción de los Índices de la *WISC-IV Spanish*, las Escalas y los Factores de la *EIWN-R PR*

<i>Índices</i> <i>WISC-IV Spanish</i> (Wechsler, 2005)	<i>EIWN-R PR</i> (Wechsler, 1992)	<i>Factores EIWN-R PR</i> (Kaufman, 1982; Carroll, Herrans y Rodríguez, 1995)
Comprensión Verbal	Escala Verbal	Comprensión Verbal
Vocabulario	Vocabulario	Vocabulario
Semejanzas	Semejanzas	Semejanzas
Comprensión	Comprensión	Comprensión
<i>Información</i>	Información	Información
---	Aritmética	---
---	<i>Retención de</i> <i>Dígitos</i>	---

VALIDACIÓN CONCURRENTE DE LA WISC-IV SPANISH

TABLA 1 (continuación)

<i>Índices</i> <i>WISC-IV Spanish</i> <i>(Wechsler, 2005)</i>	<i>EIWN-R PR</i> <i>(Wechsler, 1992)</i>	<i>Factores EIWN-R PR</i> <i>(Kaufman, 1982; Carroll,</i> <i>Herrans y Rodríguez,</i> <i>1995)</i>
Razonamiento Perceptual	Escala Ejecución	Organización Perceptual
Diseño de Bloques	Diseño de Bloques	Diseño de Bloques
---	Arreglo de Dibujos	Arreglo de Dibujos
<i>Figuras Incompletas</i>	Figuras Incompletas	Figuras Incompletas
Razonamiento Matriz	---	---
Concepto de Láminas	---	---
---	Composición de Objetos	Composición de Objetos
---	Claves	---
---	<i>Laberintos</i>	---
Memoria de Trabajo	---	Libertad de Distracción
Retención de Dígitos	---	Retención de Dígitos
---	---	Claves
Secuencia de Palabras y Números	---	---
<i>Aritmética</i>	---	Aritmética
Velocidad de Procesamiento	---	---
Claves	---	---
Búsqueda de Símbolos	---	---
<i>Cancelación</i>	---	---
Cociente Intelectual Total	Escala Total	---

Nota: Señalamos las subpruebas suplementarias en letra itálica.

Procedimiento

Luego de obtener la aprobación de la Junta de Revisión Institucional (IRB, por sus siglas en inglés) de la Universidad Interamericana de Puerto Rico y la autorización de la directora escolar, procedimos a enviar a los padres y madres de los niños de sexto, séptimo y octavo grado una carta de invitación que incluía la hoja de consentimiento informado. Los potenciales participantes depositaron el consentimiento informado debidamente cumplimentado en sobre sellado, en un buzón colocado para este propósito en la biblioteca de la escuela.

La investigadora principal administró los dos instrumentos de forma contrabalanceada a los niños participantes, luego de obtener su asentimiento. Administramos las pruebas con un intervalo de 28 a 44 días entre la primera y segunda administración, siguiendo los parámetros establecidos en la literatura científica (Knatowitz, Rediger, & Elmes, 2009; Sattler & Dumont, 2004; Wechsler, 2005).

El efecto del orden de la administración de los instrumentos y el efecto de práctica fue controlado, contrabalanceando la administración de las pruebas. Ambos instrumentos fueron administrados a los dos grupos. De manera aleatoria, se le administró primero la *WISC-IV Spanish*, seguida de la *EIWN-R PR*, a la mitad de los participantes. Para la otra mitad del grupo se invirtió el orden de la administración.

Diseño de investigación

Utilizamos el diseño intragrupos contrabalanceado. Realizamos un análisis de correlación comparando las puntuaciones estándares y a escala, respectivamente, de los Índices, Escalas, Factores y subpruebas de los dos instrumentos. También, realizamos un análisis de prueba *t* para muestras pareadas con el propósito de comparar el promedio de las puntuaciones estándar obtenidas en la Escala Total de la *EIWN-R PR* y el *FSIQ* de la *WISC IV Spanish*.

Resultados

Realizamos análisis de correlación producto-momento de *Pearson* entre los Índices de la *WISC-IV Spanish*, las Escalas de la *EIWN-R PR*, los Factores de la *EIWN-R PR* y las subpruebas de ambos instrumentos. Utilizamos las puntuaciones a escala y estándar de

VALIDACIÓN CONCURRENTE DE LA WISC-IV SPANISH

ambos instrumentos, obtenidas de los respectivos Manuales. Las correlaciones fueron corregidas por restricción en el alcance (véase Tabla 2).

Evaluamos la relación entre el *Cociente Intelectual Total* de la *WISC-IV Spanish (FSIQ)* y el *Cociente Intelectual Total* de la *EIWN-R*

TABLA 2
Coeficientes de Correlación entre las Puntuaciones Obtenidas en las Subpruebas y Escalas de la *EIWN-R PR* y Subpruebas e Índices de la *WISC-IV Spanish*

<i>EIWN - R PR</i>	<i>WISC-IV Spanish</i>			
	<i>r</i> ^a	<i>r</i> ^o	<i>FSIQ</i>	<i>VCI PRI VCI PRI</i>
Información	.72**	.72**		
Semejanzas	.58**	.58**		
Aritmética	.60**	.60**		
Vocabulario	.59**	.63**		
Comprensión	.60**	.60**		
Retención de Dígitos	.29	.33*		
Figuras Incompletas	.47**	.50**		
Diseño de Bloques	.75**	.77**		
Claves	.47**	.59**		
Escala Total	.62**		.63**	
Escala Verbal	.75**			.77**
Escala de Ejecución	.42**			.45**
Factor de Comprensión Verbal	.80**			.81**
Factor de Organización Perceptual	.50**			.51**

Nota: Todas las correlaciones presentadas en la tabla alcanzaron un nivel de significancia de .01, excepto la correlación de la subprueba de Retención de Dígitos que alcanzó un nivel de significancia de .05, al ser corregida por la restricción en el alcance.

r^a = Correlación obtenida originalmente; *r*^o = Correlación obtenida con la corrección por la restricción en el alcance; *FSIQ* = Cociente Intelectual Total; *VCI* = Índice de Comprensión Verbal; *PRI* = Índice de Razonamiento Perceptual.

PR (CIT). El resultado del análisis de correlación producto-momento de *Pearson* entre el *FSIQ* y el *CIT* demostró una correlación significativa y moderada ($r = .63, p < .01$).

Realizamos coeficientes de correlación *Pearson* para evaluar la relación entre el *Índice de Comprensión Verbal de la WISC-IV Spanish (VCI)* y la *Escala Verbal (EV) de la EIWN-R PR*. El análisis de correlación producto-momento de *Pearson* entre el *VCI* de la *WISC-IV Spanish* y la *EV de la EIWN-R PR* arrojó una correlación significativa y moderada entre las Escalas ($r = .77, p < .01$).

Evaluamos la relación entre el *Índice de Razonamiento Perceptual (PRI) de la WISC-IV Spanish* y la *Escala de Ejecución (EE) de la EIWN-R PR*. De este análisis se obtuvo una correlación significativa, aunque baja ($r = .45, p < .01$).

La evaluación de la relación entre el *Índice de Comprensión Verbal de la WISC-IV Spanish (VCI)* y el *Factor de Comprensión Verbal de la EIWN-R PR (FCV)* mostró una correlación significativa y alta entre estas variables ($r = .81, p < .01$).

Evaluamos la relación entre el *Índice de Razonamiento Perceptual de la WISC-IV Spanish (PRI)* y el *Factor de Organización Perceptual de la EIWN-R PR (FOP)*. El análisis de correlación fue significativo y moderado ($r = .51, p < .01$).

Evaluamos la relación entre las subpruebas que componen la batería estándar de la *WISC-IV Spanish* y las nueve subpruebas equivalentes de la *EIWN-R PR*. Los análisis de correlación *Pearson* realizados, alcanzaron índices entre .77 y .32 (véase Tabla 3). Se obtuvieron correlaciones significativas y con nivel moderado⁴ en casi todas las subpruebas de ambos instrumentos (*Diseño con Bloques* $r = .77, p < .01$; *Información* $r = .72, p < .01$; *Vocabulario* $r = .63, p < .01$; *Aritmética* $r = .60, p < .01$; *Comprensión* $r = .60, p < .01$; *Claves* $r = .59, p < .01$; *Semejanzas* $r = .58, p < .01$; *Figuras Incompletas* $r = .50, p < .01$). La correlación más baja fue entre las subpruebas de *Retención de Dígitos* de ambos instrumentos ($r = .33, p < .05$).

Ejecutamos un análisis de prueba *t* para muestras pareadas, para evaluar si existían diferencias entre el promedio de la Escala Total (ET) de la *EIWN-R PR* y el promedio del *Cociente Intelectual Total* de la

⁴ Cohen (1988) sugiere que una correlación alta perfecta es una correlación mayor de .80; una correlación moderada es entre .50 y .79; y una correlación baja es entre .20 y .49.

TABLA 3
Correlaciones entre las Subpruebas de la *EIWN-R PR* y la *WISC-IV Spanish*

<i>EIWN-R PR</i>	<i>Subpruebas WISC-IV Spanish</i>								
	<i>IN</i>	<i>SI</i>	<i>AR</i>	<i>VC</i>	<i>CO</i>	<i>DS</i>	<i>PCm</i>	<i>BD</i>	<i>CD</i>
IN	.72**	.65**	.24	.61**	.47**	.09	.32*	.28	.05
SI	.22	.58**	-.19	.33*	.32*	-.02	.23	.11	-.19
AR	.28	.44**	.60**	.38*	.08	.19	.43**	.25	.06
VC	.47**	.60**	.18	.63**	.43**	-.02	.12	.06	-.04
CO	.30	.42**	.03	.58**	.60**	-.05	.28	.34*	-.06
DS	.15	.10	.20	.08	.03	.33*	-.08	-.14	.19
PCm	.20	.14	.02	.01	-.04	-.08	.50**	.32*	.01
BD	.36*	.34*	.29	.36*	.13	.12	.45**	.77**	.27
CD	.24	.32*	.38*	.14	.26	.21	-.02	.20	.59**

Nota: *IN* = Información; *SI* = Semejanzas; *AR* = Aritmética; *VC* = Vocabulario; *CO* = Comprensión; *DS* = Retención de Dígitos; *PCm* = Figuras Incompletas; *BD* = Diseño de Bloques; *CD* = Claves. * $p < .05$; ** $p < .01$.

WISC-IV Spanish (FSIQ). Los resultados fueron significativos, $t = 9.15$, $p < .01$. El promedio para la *ET* de la *EIWN-R PR* fue significativamente mayor ($M = 106.48$, $SD = 12.30$) que el de la *FSIQ* de la *WISC-IV Spanish* ($M = 92.43$, $SD = 8.03$). La diferencia entre las dos puntuaciones estandarizadas de las escalas totales fue de 14 puntos.

Los resultados sugieren que las pruebas generan CI significativamente diferentes en esta muestra (véase Tabla 4).

TABLA 4
Diferencias entre las Puntuaciones de Escala Total de la *WISC-IV Spanish* y la *EIWN-R PR*

	<i>WISC-IV S</i>	<i>EIWN-R PR</i>	<i>Diferencia</i>	<i>t</i>	<i>P</i>
FSIQ	92.43	106.48	14.05	9.15	<.01

Discusión

Los resultados de esta investigación en torno a la validez concurrente de la *WISC-IV Spanish*, utilizando la *EIWN-R PR* como un criterio establecido para la evaluación del funcionamiento intelectual en niños, revelan correlaciones moderadas entre los dos instrumentos. Todas las correlaciones resultaron ser estadísticamente significativas, aunque con variaciones de intensidad entre altas, moderadas y bajas a través de los diferentes componentes.

Las correlaciones significativas más altas fueron aquellas entre el *Índice de Comprensión Verbal (VCI)* de la *WISC-IV Spanish* y el *Factor de Comprensión Verbal (FCV)* de la *EIWN-R PR* ($r = .81, p < .01$) y entre el *VCI* de la *WISC-IV Spanish* y la *Escala Verbal (EV)* de la *EIWN-R PR*. ($r = .77, p < .01$). Estos resultados son congruentes con investigaciones previas en las cuales se encontró las más altas correlaciones entre las Escalas Verbales del *WISC-III* y *WISC-R* (Lyon, 1995; Sevier *et al.*, 1994).

Como era de esperarse, el *Índice de Comprensión Verbal (VCI)* de la *WISC-IV Spanish* se correlaciona más intensamente con el *Factor de Comprensión Verbal (FCV)* y con la *Escala Verbal (EV)* de la *EIWN-R PR*, ya que es el Índice que menos cambios ha sufrido en relación al instrumento predecesor en español. El *VCI* de la *WISC-IV Spanish* consta de cuatro subpruebas (*Semejanzas, Vocabulario, Comprensión e Información*), todas presentes en la *EV* y el *FCV* de la *EIWN-R PR*. Esta congruencia en subpruebas supone similitudes en el peso lingüístico y de conocimiento cristalizado compartidos entre el *VCI*, el *FCV* y la *EV* de ambos instrumentos.

La correlación más baja obtenida fue entre el *Índice de Razonamiento Perceptual (PRI)* de la *WISC-IV Spanish* y la *Escala de Ejecución (EE)* de la *EIWN-R PR* ($r = .45, p < .01$). En el ámbito de la evaluación psicológica no se ha documentado previamente esta relación de baja intensidad entre el *PRI* y la *EE*. Sin embargo, se han encontrado estudios que señalan las más bajas correlaciones entre las Escalas de Ejecución de la *WISC-III* y la *WISC-R* (Lyon 1995; Sevier *et. al.*, 1994.)

En esta investigación, la baja correlación obtenida entre el *PRI* y la *EE* puede estar asociada a la falta de contenido común a las dos Escalas. De las seis subpruebas que componen la *EE* de la *EIWN-R PR*, una pasó a ser suplementaria (*Figuras Incompletas*) y tres fueron suprimidas por completo (*Arreglo de Dibujos*, *Composición de Objetos* y *Laberintos*) del *WISC-IV Spanish*, para reducir el énfasis en rapidez y límites de tiempo. Por lo tanto, la *Escala de Ejecución* de la *EIWN-R PR (EE)* y el *Índice de Razonamiento Perceptual* de la *WISC-IV Spanish (PRI)* sólo comparten una subprueba (*Diseño con Bloques*). Otra diferencia marcada es que el *PRI* de la *WISC-IV Spanish* incorpora dos nuevas subpruebas (*Concepto de Láminas* y *Razonamiento Matriz*), con el propósito de incrementar las medidas de razonamiento fluido. Todos estos cambios hacen que el *PRI* y la *EE* no sean tan comparables o equivalentes, ya que aparentemente evalúan factores cognoscitivos diferentes.

La correlación moderada ($r = .62$) obtenida entre el *FSIQ* del *WISC-IV Spanish* y la *Escala Total (ET)* del *EIWN-R PR* era de esperarse en vista de las diferencias sustanciales entre los dos instrumentos. Por un lado, el *FSIQ* del *WISC-IV Spanish* está basado en cuatro Índices, los cuales proveen una representación más amplia de funcionamiento cognoscitivo que la *ET* del *EIWN-R PR*, el cual está basado en dos Escalas (*EV* y *EE*). La literatura reporta resultados mixtos, indicando correlaciones significativas, pero variables entre el *WISC-III* y el *FSIQ* del *WISC-R* que fluctúan entre .85 y .66 (Lyon 1995; Sevier *et. al.*, 1994), los cuales son dos instrumentos más parecidos entre sí. Por lo tanto, la correlación significativa y moderada obtenida entre el *FSIQ* del *WISC-IV Spanish* y la *ET* del *EIWN-R PR*, a pesar de ser éstos los dos instrumentos más disímiles, se convierte en un hallazgo interesante.

Según se esperaba, las correlaciones más altas obtenidas fueron entre las subpruebas similares del instrumento predecesor. La única

excepción a este patrón fue la baja correlación obtenida entre la subprueba de *Retención de Dígitos* de ambos instrumentos. A pesar de que la subprueba ha cambiado muy poco de una versión a otra, este resultado aparenta estar asociado a elementos normativos, muestreo y puntuaciones estandarizadas que pudieron afectar el proceso de evaluación.

El interés secundario de los investigadores fue conocer si las puntuaciones totales obtenidas en el *WISC-IV Spanish* son comparables con las de la *EIWN-R PR*. Los resultados de esta investigación fueron en la dirección esperada y son consistentes con los hallazgos de Kanaya *et al.* (2003), quienes al comparar la *WISC-R* y la *WISC-III* encontraron CI aproximadamente seis puntos más bajos en la versión más reciente del instrumento. Al igual que la investigación de Kanaya y sus asociados (2003), la actual investigación reveló puntuaciones significativamente más bajas en la recientemente publicada *WISC-IV Spanish*. En este caso, se obtuvieron CI 14 puntos menos en la *WISC-IV Spanish*. Parte de la discrepancia (entre cinco a nueve puntos, aproximadamente) era de esperarse y puede atribuirse al efecto Flynn, ya que hay 14 años de diferencia entre la publicación de la *EIWN-R PR* en 1992 y el año en que se efectuó esta investigación.

Otros tres factores que pudieron contribuir a la diferencia en puntuaciones entre las Escalas Totales son: el error estándar de la medida, los cambios normativos y el tamaño de la muestra. Hay que considerar el error estándar de la medida, ya que las puntuaciones son un estimado de la puntuación real, es decir, que las puntuaciones obtenidas están dentro de un intervalo de confianza donde se encuentra la puntuación real. Por otro lado, la diferencia en puntuaciones puede ser explicada por variables normativas asociadas al nuevo instrumento. Además, las características de la actual muestra y la de normalización son diferentes. En particular, la muestra que se utilizó era pequeña y homogénea, variables que pueden provocar mayor variabilidad y error, finalmente redundando en puntuaciones menos predecibles.

Aunque es difícil saber con certeza por qué los actuales hallazgos revelan mayor disparidad entre las puntuaciones totales que aquellos reportados en la literatura científica, es posible que la discrepancia observada esté relacionada al hecho de que la literatura reseñada primordialmente evaluó la *WISC-R* y la *WISC-III*, versiones más cercanas y parecidas entre sí. En la actual investigación, se compararon dos instrumentos más alejados en términos de fecha de publicación y más

disímiles en esquema conceptual y teórico. Este hallazgo amerita explorarse con mayor detenimiento.

Aunque esta investigación provee evidencia de una relación moderada entre los dos instrumentos, el uso de la recientemente publicada *WISC-IV Spanish* ha de tomarse con extrema cautela en vista de las diferencias encontradas entre sus puntuaciones y aquellas obtenidas con la *EIWN-R PR*. Además, la ausencia de normas para los niños en Puerto Rico constituye una seria limitación para su utilización en la Isla.

En particular, preocupa la posible utilización de la *WISC-IV Spanish* en Puerto Rico, en procesos de reevaluación y de toma de decisiones educativas. Por ejemplo, muchos niños que en el futuro inmediato podrían ser reevaluados con la *WISC-IV Spanish*, probablemente cuentan también con evaluaciones iniciales con la *EIWN-R PR*. Esta diferencia en puntuaciones totales puede tener un impacto en la elegibilidad para servicios de niños que sean evaluados con el instrumento normalizado más recientemente, ya que podrían puntuar significativamente por debajo de las puntuaciones obtenidas en la evaluación inicial. Esta preocupación ha sido planteada en un estudio que se llevó a cabo con niños que habían sido previamente diagnosticados con retraso mental y luego fueron reevaluados con la versión más reciente del mismo instrumento (Kanaya *et al.*, 2003).

De acuerdo con Kanaya *et al.* (2003) y Flynn *et al.* (2007), la precaución en torno a la utilización de un nuevo instrumento en procesos de reevaluación debe ser mayor durante los primeros cinco años de la publicación del mismo. Según estos investigadores, ése es el lapso de tiempo en que es más probable que las puntuaciones de la prueba nueva sean comparadas con las puntuaciones “infladas” de la prueba antecesora. Por esta razón, se sugiere extrema cautela al utilizar la *WISC-IV Spanish*, en particular al realizar reevaluaciones, al emitir recomendaciones y al tomar decisiones clínicas, legales o sociales, utilizando como base los hallazgos obtenidos con ese instrumento nuevo y carente de normas locales.

Las limitaciones de la actual investigación deben ser consideradas cuando se interpretan los resultados. La limitación principal fue la selección de una muestra homogénea escogida por disponibilidad. Por consiguiente, los hallazgos de esta investigación sólo evidencian la validez concurrente de la *WISC-IV Spanish* para la muestra evaluada. Otra de las limitaciones fue el tamaño de la muestra. La misma fue

relativamente pequeña ($n = 40$) y limitada a un sólo sector geográfico de la Isla. El limitado tamaño y la limitada extensión geográfica de la muestra no permiten hacer juicios abarcadores y otros análisis que permitan identificar con mayor claridad la aplicabilidad del nuevo instrumento.

En futuras investigaciones, recomendamos incorporar participantes cuyas edades fluctúen entre seis a 16 años, edades para las cuales están normalizados ambos instrumentos de interés. Además, estudios futuros deben examinar la relación entre el *WISC-IV Spanish* y la *EIWN-R PR* en una muestra heterogénea, que incluya estudiantes de escuela pública y privada con diversos niveles socioeconómicos. Se recomienda que se aumente el tamaño de la muestra y se lleven a cabo investigaciones similares en otras regiones geográficas de la Isla.

Esta investigación también suscita varias interrogantes que se podrían evaluar en futuros esfuerzos: ¿En una muestra representativa, se observará la diferencia significativa entre el *FSIQ* de la *WISC-IV Spanish* y la Escala Total de la *EIWN-R PR*? ¿Qué resultados se obtendrán utilizando una muestra de estudiantes con necesidades especiales? ¿Qué patrones de puntuación serán típicos en diversos grupos de niños puertorriqueños?

La tarea de desarrollar una base de datos adecuada, en torno a las propiedades psicométricas de instrumentos que están al alcance de los psicólogos en Puerto Rico, es muy importante. Se espera que esta investigación haya contribuido en esa dirección.

Como tema a ponderar y dialogar queda la forma en que los psicólogos e investigadores en el campo de la medición psicológica en Puerto Rico responderán a las exigencias sociales y legales en torno a dos asuntos de importancia. Por un lado, está la respuesta que se proveerá ante la creciente necesidad de instrumentos válidos, confiables y con normas locales recientes. Como segundo asunto, se encuentra cómo lograr que las medidas desarrolladas para niños y adolescentes en Puerto Rico incorporen los avances teóricos y metodológicos estipulados por ley, los cuales atañen a la evaluación de los múltiples procesos cognoscitivos y de procesamiento de información. Éstos son grandes retos, particularmente en el contexto de las limitaciones económicas y comerciales actuales que no necesariamente favorecen estos desarrollos en la Isla.

REFERENCIAS

- Baddeley, A. (2003). Working memory: Looking back and looking forward. *Natural Reviews Neuroscience*, 4, 829-839.
- Burns, T. & O'Leary, S. D. (2004). Wechsler Intelligence Scale for Children-IV: Test review. *Applied Neuropsychology*, 4, 235-238.
- Carroll, J. F., Herrans L. L., & Rodríguez, J. M. (1995). Análisis factorial de la EIWN-R de Puerto Rico con niños de 11 niveles de edad, entre los 6 y 16 años. *Revista Latinoamericana de Psicología*, 27, 2, 187-206.
- Carlton, M. & Sapp, G. (1997). Comparison of WISC-R and WISC-III scores of urban exceptional students. *Psychological Reports*, 80, 3, 755-760.
- Cohen, J. (1988). Statistical power analysis for the behavioral sciences. (2nd. ed.) Hillsdale, NJ: Erlbaum.
- Flanagan, D. P. & Kaufman, A. S. (2004). *Essentials of WISC-IV assessment*. New Jersey: John Wiley & Son.
- Flynn, J. (1987). Massive IQ gains in 14 nations: What IQ test really measure. *Psychological Bulletin*, 10, 2, 171-191.
- Flynn, J. R. (1994). IQ gains over time. En R. J. Sternberg (Ed.), *Encyclopedia of human intelligence* (pp. 617-623). Nueva York, Nueva York: Macmillan.
- Flynn, J. R. (1999). Searching for justice: The discovery of I.Q. gains over time. *American Psychologist*, 54, 5-20.
- Flynn, J. & Weiss, L. (2007). American IQ gains from 1932 to 2002: The WISC subtest and educational progress. *International Journal of Testing*, 7, 2, 209-224.
- Herrans, L. L. & Rodríguez, J. M. (1992). *Informe final proyecto EIWN-R de Puerto Rico*, Departamento de Salud, Secretaría Auxiliar de Salud Mental, San Juan, Puerto Rico.
- Individuals with Disabilities Education Improvement Act of 2004 (IDEIA), Pub. L. No.108-466, 118 Stat. 2647 (2004). [Amending 20 U.S.C. §§ 1400 et seq.].

- Jones, P. (1996). Dissertation Abstracts International Section A: Humanities and Social Sciences. [Electronic version]. *US ProQuest Information & Learning*, 56, 10, 4041.
- Kanaya, T., Scullin, M., & Ceci, S. (2003). The impact of rising IQ scores on American society via mental retardation diagnoses. *American Psychologist*, 58, 10, 778-790.
- Kaplan, E. (1988). Clinical neuropsychology and brain function: Research, measurement and practice: A process approach to neuropsychological assessment. En T. Boll & B. K. Bryant (Eds.), *Clinical neuropsychology and brain function: research, measurement, and practice* (pp. 129-167). Washington, DC: American Psychological Association.
- Kaufman, A. (1975). Factor analysis of the WISC-R at 11 age-levels between 6 ½ and 16 ½ years. *Journal of Consulting and Clinical Psychology*, 43, 2, 135-147.
- Kaufman, A. (1982). *Psicometría razonada con el WISC-R*. Nueva Jersey: John Wiley and Sons.
- Knatowitz, B. H., Rediger, H. L., & Elmes, D. (2009). *Experimental psychology*. (9th ed.). Belmont, USA; Wadsworth.
- Lyon, M. A. (1995). A comparison between WISC III and WISC-R scores for learning disabilities reevaluations. *Journal of Learning Disabilities*, 28, 4, 253.
- Matlin, M. W. (2005). *Cognition*. Nueva Jersey: John Wiley & Son.
- Morris, C. G. & Maisto A. (2001). *Psicología*. México: Pearson Education.
- Prifitera, A., Saklofske, D. H., & Weiss, L. G., (2005). *WISC-IV clinical use and interpretation: Scientist-practitioner perspectives*. San Diego, CA: Elsevier
- Pons, J. (2007). Uso en Puerto Rico de la Escala de Inteligencia WISC-IV- Español. *Boletín de la Asociación de Psicólogos de Puerto Rico*, 1, 6-8.
- Sattler, J. M. & Dumont, R. (2004). *Assessment of children WISC-IV and WPPSI-III supplement*. California: Jerome M. Sattler, Publisher.

VALIDACIÓN CONCURRENTE DE LA WISC-IV SPANISH

- Sevier, R. C. & Bain, S. K. (1994). Comparison of WISC-R and WISC III for gifted students. *Roeper Review*, 17, 1, 39.
- Wechsler, D. (1974). *Wechsler Intelligence Scale for Children–Revised*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (1991). *Wechsler Intelligence Scale for Children–Third Edition*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (1992). *Manual de la Escala de Inteligencia Wechsler para Niños–Revisada de Puerto Rico*. San Antonio, Texas: The Psychological Corporation.
- Wechsler, D. (2003). *Wechsler Intelligence Scale for Children–Fourth Edition*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (2005). *Wechsler Intelligence Scale for Children–Fourth Edition, Spanish Manual*. San Antonio, Texas: The Psychological Corporation.
- Weiss, L. G., Saklofske, D. H., Prifitera, A., & Holdnack, J. A. (2006). *WISC-IV advanced clinical interpretation*. San Diego, CA: Elsevier.