

ADAPTACION Y NORMACION DE LA PRUEBA DE VOCABULARIO PEABODY EN UNA MUESTRA DE SUJETOS CUBANOS

Armando Piñeiro*; Mayra Manzano**; Gonzalo Inguanzo*; Vivian Reigosa*; Aymara Morales*** y Caridad Fernández***

* Centro de Neurociencias de Cuba, La Habana

** Facultad de Psicología de la Universidad de La Habana

*** Escuela Especial "Paquito Rosales

RESUMEN

En este trabajo se muestran los resultados de la adaptación y normación de la Prueba de Vocabulario Peabody, (Dunn, 1959) en una muestra de sujetos cubanos. Inicialmente se tradujo el material verbal de la versión norteamericana de la prueba y se modificaron algunas láminas. Posteriormente la prueba fue aplicada a 775 sujetos sanos de entre 2½ y 17½ años. Se calcularon los valores normativos por edad y las normas percentiles con sus respectivos intervalos confidenciales. La correlación entre la edad cronológica y el puntaje obtenido en la prueba resultó muy significativo ($r=.91$ $p=0.000$). Se determinó para cada grupo de edad la palabra de inicio de la prueba. Se calcularon los coeficientes de confiabilidad y los errores de medida para cada grupo de edad, utilizando el método split-half. Todos los coeficientes de confiabilidad resultaron estadísticamente significativos ($r=.60$ ó mayor $p<.05$). El coeficiente de confiabilidad calculado con la técnica test-retest también fue significativo ($r=.72$ $p=.000$). Los dos estudios de validez realizados evidencian que la prueba evalúa eficientemente el vocabulario.

ABSTRACT

This paper shows the result of the Cuban adaptation and standardization of the Picture Peabody Vocabulary Test (Dunn, 1959). The verbal material of the American version of the test was translated and some pictures were modified. The test was applied to 775 subjects between 2½ and 17½ years. The normative data for age and the percentiles with their respective confidential intervals were calculated. The correlation between the Chronological age and the result was very significant ($r=.91$ $p=0.000$). The first item of the test was determined for each group of age. The reliability calculated using the split-half method was significant ($r>.60$ $p<0.05$). The reliability calculated using test-retest also was significant ($r=.72$ $p=0.000$). The validity of the Cuban adaptation of the test was significant.

La Prueba de Vocabulario Peabody (PVP), desarrollada por L. Dunn, (1959), es uno de los instrumentos más utilizados actualmente para evaluar el vocabulario. Las ventajas más significativas de la PVP están dadas por lo atractivo que resulta para la mayoría de los sujetos, lo que condiciona un buen rapport, no precisa de una preparación especializada para su administración, la aplicación no consume mucho tiempo (15-20 minutos), no precisa de una respuesta oral y el rango de edad que cubre es bastante amplio.

Según su autor la prueba está diseñada para medir fundamentalmente el vocabulario receptivo, además de la inteligencia verbal o aptitudes escolares entre los 2 y los 17 años de edad.

La prueba consiste en 150 palabras y 150 placas que contienen cada una cuatro láminas (objetos manufacturados, animales, acciones humanas,

plantas, etc.). El sujeto debe seleccionar para cada palabra la lámina de la placa que más se le parezca.

No se aplican todas las palabras de la prueba, sino sólo aquellas que están en torno a la edad cronológica del sujeto, no se aplican ni las palabras muy fáciles que son propias de edades muy pequeñas, ni las muy difíciles propias de edades superiores. Para cada edad se ha determinado una palabra de inicio. A partir de la palabra de inicio se determina el ítem base, todos los ítems que estén por debajo del ítem base se consideran que pueden ser respondidos correctamente. También se determina el ítem techo por encima del cual se considera que el sujeto no puede responder con éxito.

Al aplicar la prueba se obtiene un puntaje bruto, el cual se calcula restando la cantidad de palabras en las que el niño fue evaluado menos la cantidad de

errores. Este puntaje bruto debe convertirse en una puntuación normativa.

En la PVP se ofrecen tres tipos de normas: por edad, percentiles y las puntuaciones típicas derivadas (CI). El evaluador puede escoger el tipo de dato normativo que desee, puesto que todos ellos son equivalentes.

La PVP fue normada para sujetos angloparlantes (Dunn y Dunn, 1981; Dunn, Dunn, Whetton y Pintillie 1982). Para el idioma español se han realizado también varios trabajos normativos (Dunn, 1985; Dunn, Padilla, Lugo y Dunn 1986). Las versiones de la prueba utilizadas para otras poblaciones hispanohablantes contienen palabras y figuras que no son familiares en Cuba.

El objetivo de este trabajo fue adaptar la PVP para que pudiera ser aplicada a sujetos cubanos y normar la prueba.

El trabajo de adaptación y normación de la PVP para sujetos cubanos constó de varias etapas:

1. La traducción de las 150 palabras de la prueba original.
2. Modificación de algunas figuras.
3. Aplicación de la prueba a una muestra de sujetos.
4. Reordenamiento de las palabras.
5. Análisis de los errores.
6. Cálculo de las normas.
7. Estimación de la confiabilidad.
8. Estimación de la validez.

1. La traducción de las 150 palabras de la prueba original

Se tradujeron al español las palabras originales de la prueba, utilizando para ello el Webster's New Collegiate Dictionary. Se siguieron una serie de criterios para seleccionar la aceptación más adecuada de la palabra original.

Puesto que en la prueba se utilizan colectivos, gerundios, adjetivos y sustantivos, se trató que la traducción respetara la función gramatical de la palabra original, por ejemplo la palabra blowing fue traducida como soplando. En aquellos casos en que fue posible se evitaron traducciones que desembocaran en una palabra compuesta. Puesto que en la prueba la dificultad de las palabras es creciente se trató de que la traducción mantuviera intacto este nivel de dificultad. No fueron traducidas aquellas palabras para las cuales en nuestro idioma el uso ha acuñado la palabra en otro idioma, por ejemplo dial y kayak.

Todos estos criterios nos permitieron una traducción que respeta, en lo posible, el espíritu de

la prueba original.

2. Modificación de algunas figuras

La versión original de la prueba contiene ilustraciones que no se ajustan a nuestra cultura. Por ello se realizaron cambios en las imágenes, respetando las características de las líneas de las figuras originales, su forma, la proporción y el tamaño.

En algunos casos los cambios se relacionaron con algún texto que aparecía en las láminas, por ejemplo en la lámina 2 el nombre del jabón; 33 la dirección y el remitente del sobre de carta. Fueron transformadas algunas figuras por representaciones de las mismas más cercanas a nuestra realidad, por ejemplo, en la placa 35 la estrella del sheriff fue transformada en una estrella de destacado. Se suprimieron en algunas figuras ciertos trazos por traer confusiones con otras: en la figura 62 se omitió el asa del embudo y en la 131 las líneas que representaban el viento.

3. Aplicación de la prueba a una muestra de sujetos

La prueba fue aplicada a una muestra de 775 sujetos cubanos, 401 hembras y 374 varones. La edad de los sujetos osciló entre los 2.06 y los 17.06 años. La edad promedio fue 9.52 años (d.s.=4.29). La muestra fue dividida en 15 grupos de al menos 50 sujetos cada uno. A continuación se muestran los intervalos de edad por grupo y la cantidad de sujetos incluidos en cada uno de ellos.

Tabla 1: Cantidad de sujetos incluidos en cada grupo de edad.

Edades	n	Edades	n
2.05-3.05	53	10.06-11.05	53
3.06-4.05	50	11.06-12.05	50
4.06-5.05	54	12.06-13.05	50
5.06-6.05	50	13.06-14.05	57
6.06-7.05	54	14.06-15.05	50
7.06-8.05	50	15.06-16.05	50
8.06-9.05	52	16.06-17.05	50
9.06-10.05	52		

Los niños preescolares fueron seleccionados en Círculos Infantiles. Los niños escolares fueron seleccionados en varias escuelas de EGB. Los adolescentes y jóvenes estudiaban en escuelas de Secundaria Básica, Preuniversitarios y en centros de enseñanza Técnica. Todos los sujetos vivían en Ciudad de La Habana y La Habana.

Se trató que los sujetos incluidos en la muestra no tuvieran alteraciones en el lenguaje, ni pérdidas auditivas. Con este propósito, en el caso de los niños, se entrevistaron a las educadoras de los círculos infantiles, a las auxiliares y a las maestras.

Los adolescentes y jóvenes fueron entrevistados por los investigadores, quienes decidieron si podían ser incluidos o no en la investigación.

Para cada sujeto se recogió la escolaridad de ambos padres. La siguiente tabla muestra la distribución de estos datos.

Tabla 2: Escolaridad de los padres de los sujetos de la muestra.

Grado escolar	Madres	Padres
Primaria sin terminar	1	18
Primaria terminada	29	27
Media (secundaria)	93	82
Media (superior)	258	229
Universitaria	207	226
Sin datos	187	193

La prueba fue aplicada por siete examinadores entrenados. Se recogió para cada sujeto la respuesta emitida ante cada una de las 150 láminas. Como se deseaba tener normas nuevas, se aplicó la prueba completa sin calcular ni el ítem base, ni el techo. Se determinó la cantidad de respuestas correctas por sujeto.

Además se anotó el número de ejemplos necesarios para apropiarse de la prueba, la calidad del rapport establecido y si el sujeto intentaba adivinar o no la posible respuesta.

4. Reordenamiento de las palabras

Por la forma de aplicación se obtuvo para cada palabra el porcentaje de respuestas correctas. A partir del porcentaje de acierto en cada palabra para cada grupo de edad, por una parte se reordenaron las palabras en dificultad creciente y por otra se calculó la palabra en que debe iniciarse la aplicación de la prueba en función de la edad de los sujetos. El reordenamiento se realizó primero automáticamente y luego se hicieron una serie de ajustes de forma manual. Para determinar el orden de las palabras, se aplicaron los siguientes criterios:

- Edad en la que comienzan a aparecer los aciertos
- Edad en la cual se alcanza el 50 % de aciertos
- El comportamiento subsiguiente (después de alcanzar el 50 % de los aciertos en las otras edades).

De ese modo si para una palabra los aciertos comienzan a los 10 años, se considera más difícil que una en la que comiencen antes, a los 8 años, pero además se observó a qué edad se alcanzó el 50 % de aciertos. Mientras se alcance a edades más tempranas el 50 % de aciertos se considera que la palabra es más fácil y, por lo tanto, debe ir ubicada

más temprano en la prueba. Por último se tuvo en cuenta cómo se comportan los porcentajes de aciertos en las edades posteriores a la edad en que alcanza el 50 % de aciertos. Si la palabra no llega a alcanzar en ninguna edad un 80 por ciento de aciertos, se supone que ella debe ocupar un lugar avanzado en el orden de la prueba.

Sólo 6 palabras no cambiaron de lugar en el nuevo orden establecido con respecto al orden original de la prueba. A continuación se muestran estas palabras y entre paréntesis su número en la prueba: carro (1), gallina (10), jamo (60), trasportación (63), dirigiendo (69) y obelisco (136).

El monto del cambio, es decir, el número de espacios hacia adelante o hacia atrás en los que se desplazó una palabra osciló entre 0 y 64.

El sentido más frecuente de cambio fue para disminuir el número de orden de la palabra:

79 cambios son de ese sentido. Por ejemplo la palabra "absorber" era la 147 en la prueba original, en el nuevo orden es la palabra 89, esta lámina fue desplazada 58 láminas hacia adelante en la prueba, 65 palabras aumentaron su número de orden, la palabra dial en la prueba original ocupaba el número 46 y en el nuevo orden es la número 110.

A continuación se muestra la cantidad de palabras que fueron desplazadas en esta adaptación de la PVP con relación a la versión original y el monto de este desplazamiento.

Tabla 3: Cantidad de palabras desplazadas en posición en relación con la versión original.

Monto del Desplazamiento	Disminución del número de orden	Aumento del número de orden
Entre 1 y 10	35	27
Entre 11 y 20	22	15
Entre 21 y 30	14	8
Más de 30	8	15

Para calcular la palabra en la que se debe comenzar a aplicar la prueba en cada edad se tomó, sobre la base del nuevo orden establecido, aquella palabra en la que los sujetos, ordenados en grupos de edad, alcanzaron un 85 % de aciertos.

Como promedio nuestros datos nos hacen desplazar la palabra de comienzo de la prueba en 17 lugares por encima con relación a la prueba original. Por ejemplo para niños comprendidos entre los 6.06 y los 7.05 en la versión original la prueba debe comenzar en la palabra 40 mientras que la versión cubana para la edad anteriormente señalada la palabra de inicio debe ser la 54, hay una diferencia de 14 palabras. En un solo grupo de edad nuestros sujetos deben comenzar en una palabra inferior a la fijada en la versión original. En la edad

comprendida entre los 5.06 años la palabra de inicio en la versión original es la 40 y en nuestra versión es la 36.

La Tabla 4 muestra para cada intervalo de edad la palabra donde se debe comenzar la aplicación de la prueba.

Tabla 4: Palabra de inicio para cada grupo de edad.

Edades	n
2.05-3.05	17
3.06-4.05	22
4.06-5.05	29
5.06-6.05	36
6.06-7.05	54
7.06-8.05	58
8.06-9.05	61
9.06-10.05	73
10.06-11.05	83
11.06-12.05	90
12.06-13.05	94
13.06-14.05	99
14.06-15.05	101
15.06-16.05	117
16.06-17.05	120

5. Análisis de los errores

El análisis de los errores cometidos por los sujetos en cada placa permitió evaluar si existía alguna consistencia en estos, indicando una posible confusión entre la lámina correcta y las restantes de la placa.

En general, en ninguna placa se encontró un patrón importante de error (50 % o más de los errores en una lámina alternativa), que determinara un cambio de las láminas de la placa. Sin embargo, se realizaron algunas sustituciones cuando:

- El error recaía sobre una lámina incorrecta en un 25 por ciento o más de las veces en toda la muestra.
- La suma de los porcentos de error de dos láminas de la placa sumaba 33 por ciento o más en toda la muestra.

- Cuando los errores fueron altos (25 por ciento o más) en un grupo de edad, aunque el porcentaje de error en toda la muestra estuviera por debajo del 25 por ciento.

En total fueron se sustituyeron 11 láminas en 11 placas por láminas de otras placas. En cada placa se cambió una sola lámina.

Por ejemplo en la placa que corresponde a la palabra dial había una lámina que representaba una tapa de alcantarilla. Sobre esta lámina recayó la respuesta de un 33 % de los sujetos. Puesto que la forma común de ambos objetos puede llevar a la confusión la lámina fue cambiada y en sustitución se colocó una pinza. En la placa correspondiente a la palabra vítreo aparecía una lámina con una copa que evidentemente era de cristal, pero además aparecía una lámina de un candelabro que dado la forma general de los dibujos (líneas simples, en blanco y negro), podía parecer que también era de cristal, de hecho 37.30 % seleccionaron esta lámina y no la correcta, por lo que se cambió esta lámina por una que representa una rueda.

6. Cálculo de las normas

Se determinaron tres tipos de normas, por edad, percentiles y CI, estas normas son comparables entre sí, el investigador puede utilizar cualquiera de ellas. Para poder calcular las normas primeramente se determinó para cada grupo de edad el tipo de distribución estadística que seguía la puntuación bruta obtenida en la prueba, utilizando la prueba de Kolmogorov-Smirnov para la distribución. En todas las edades los datos tuvieron una distribución normal ($p > .05$)

Para cada edad se calcularon los valores de la media y la desviación standard, los valores de puntuación mínimo y máximo. Se calcularon también los percentiles 10, 25, 50, 75 y 90.

La siguiente tabla muestra para cada grupo de edad la estadística descriptiva del puntaje en la prueba.

Tabla 5: Estadística descriptiva de las puntuaciones por grupo de edad.

Edades	Media	d.s.	Rango	10 %il	25 %il	50 %il	75 %il	90 %il
2.06-3.05	28.00	8.03	18-50	19	21	26	33	38
3.06-4.05	41.22	10.05	25-61	29	33	41	49	56
4.06-5.05	53.37	11.19	28-80	39	48	53	59	68
5.06-6.05	70.30	14.80	40-110	52	60	70	79	90
6.06-7.05	80.22	14.06	51-107	61	68	80	92	98
7.06-8.05	86.12	10.89	62-114	73	78	84	94	100
8.06-9.05	94.20	11.23	72-118	77	88	95	102	106
9.06-10.05	96.50	11.98	72-118	77	90	99	105	108
10.06-11.05	108.26	14.54	70-137	93	102	109	117	125
11.06-12.05	113.62	12.82	82-146	97	104	116	120	129
12.06-13.05	121.16	11.71	79-136	109	114	125	130	132
13.06-14.05	123.61	10.14	98-145	109	118	125	130	136
14.06-15.05	123.72	8.13	109-141	115	116	124	129	135
15.06-16.05	130.84	10.19	99-143	116	125	134	137	141
16.06-17.06	137.54	4.63	129-148	132	135	137	140	145

Se observa un aumento del puntaje con la edad, al igual que en la versión original.

Al comparar los puntajes de la muestra normativa cubana y la norteamericana (Forma A) para las edades comprendidas entre 5.06-6.05 y 16.06-17.05, se observa que los sujetos cubanos obtienen como promedio 25 puntos más que los norteamericanos (d.s.=5.30). Esta diferencia podría tener en su base las diferencias idiomáticas, que se mantienen aunque se haya tratado de que la traducción de las palabras al español tuvieran el mismo grado de dificultad que las palabras inglesas de la versión original.

Se destaca por otra parte que todos los sujetos de la muestra cubana tienen puntajes superiores a 0, el puntaje más pequeño fue 18, encontrado en el primer grupo. En ningún grupo de edad se alcanza el valor techo (150), el sujeto que mayor puntaje alcanzó en la prueba logró 148 puntos y pertenecía al último grupo. La diferencia entre el techo y la media del grupo de sujetos mayores fue de alrededor de 12 puntos.

Las desviaciones standard son pequeñas (entre 4.63-14.80 puntos), esperadas para este tipo de prueba (en Dunn, 1965 las desviaciones standard para la Forma A oscilaron entre 7.52 y 14.97).

El puntaje se incrementó como promedio de una edad a otra en 8.5 puntos (ds=6.6, mediana=7.5). El máximo incremento del puntaje se produjo entre los grupos de edad 4.06-5.05 y 5.06-6.05 (incremento=27 puntos). El incremento mínimo ocurrió entre las edades 13.06-14.05 y 14.06-15.05 (incremento=0.11 puntos).

Para determinar si las diferencias observadas entre los puntajes de las distintas edades eran estadísticamente significativas se calculó un análisis de varianza (ANOVA One-Way). Los resultados del análisis de varianza confirman que las diferencias de los puntajes son significativas: $F(14,760)=462.55$ $p<0.0001$.

Se compararon las medias de los puntajes edad por edad usando la prueba de Newman-Keuls. Las medias resultaron significativamente diferentes en la mayoría de los contrastes ($p<0.05$). Cuatro comparaciones no resultaron estadísticamente significativas, estas fueron: entre 8.06-9.05 y 9.06-10.05 (incremento=2 puntos), entre 12.06-13.05 y 14.06-15.05 (incremento=2 puntos) y entre 13.06-14.05 y 14.06-15.05 (incremento=0.11 puntos).

La media del puntaje de las niñas fue 94.2 (ds=34.5) y la media del puntaje obtenido por los varones fue 93.10 (d.s.=34.4). Estas diferencias no fueron significativas estadísticamente al realizar una comparación utilizando el estadígrafo t de Student $t(773)=.63$ $p>0.10$.

Al correlacionar la edad de los sujetos y el puntaje obtenido por cada uno de ellos se obtuvo un

coeficiente de correlación alto y positivo $r=0.918$ $p=0.000$.

Se calculó la ecuación de regresión del puntaje a partir de la edad. Esta ecuación permite predecir para cualquier edad la cantidad de puntos que debe obtener un sujeto en la prueba incluso si está fuera del intervalo de edades para los que se hizo el trabajo de normación. La ecuación de regresión obtenida fue:

$$\text{puntaje}=23.454+7.37* \text{ edad cronológica.}$$

Sobre la base de la relación encontrada entre la edad y el puntaje se calcularon las normas de edad, que permiten determinar la edad de vocabulario de un sujeto a partir de su puntuación. Por ejemplo, nuestras normas de edad establecen que un niño que alcance un puntaje de 83 tiene una edad de vocabulario de 7 años, mientras que un niño que alcance una puntuación de 47 puntos tiene una edad de vocabulario de 4 años. A partir de los datos obtenidos en el estudio de la confiabilidad (que se reseña más adelante), se estimaron los intervalos de confianza para las normas de edad.

Otro tipo de norma que se calculó fueron las percentiles. Para cada grupo de edad, se obtuvo el percentil asociado a cada puntaje. Por ejemplo el puntaje 70 se corresponde para el intervalo de edad de 4.06-5.05 con el 95 percentil, en el intervalo de edad 5.06-6.05 con el 50 percentil y en el grupo de edad 6.06-7.05 con el 25 percentil.

A partir de los datos normativos obtenidos para cada grupo de edad es posible calcular el C.I., para ello es necesario aplicar la siguiente fórmula:
 $C.I.=100+(15/d.s.)*(P.S.-MPGN)$

Donde: C.I.=Coeficiente de Inteligencia, d.s.=desviación standard del grupo de edad normativo del sujeto a evaluar, P.S.=puntaje del sujeto, MPGN=media de puntaje del grupo normativo del sujeto a evaluar.

Por ejemplo si se desea conocer el CI obtenido por un niño de 7 años cuyo puntaje en la prueba fue de 84, se aplica la fórmula de la siguiente manera:

$$C.I.=100+(15/14.06)*(84-80.22)$$

Los datos del grupo normativo se extraen de la Tabla 2

$$C.I.=104$$

7. Estimación de la confiabilidad

Las pruebas psicométricas, al igual que otros tipos de medida, en la evaluación tienen algún grado de error. Las causas de estos errores pueden ser diversas, por ejemplo deficiencias en la construcción del test, problemas en la calificación, variación de estilos de aplicación de los examinadores, fluctuaciones en el rendimiento de los sujetos, etc. La confiabilidad puede ser definida como la parte de

la varianza de la prueba libre de errores, e indica hasta qué punto son replicables los resultados que obtiene un sujeto en una prueba.

Para determinar si la adaptación cubana de la PVP era confiable se utilizaron dos procedimientos diferentes. Para determinar la consistencia interna se correlacionaron los puntajes brutos de los ítems impares con los puntajes brutos de los ítems pares obtenidos por los sujetos de la muestra normativa dividida por grupo de edades. Para ello se utilizó el coeficiente de correlación de Pearson y luego se aplicó la corrección de Sperman-Brown (Guilford, 1950). A partir de los coeficientes de correlación se calculó para cada grupo de edad los errores de medida.

Para ellos se utilizó la fórmula:

$$EM = \pm \sqrt{1 - r_{pi}}$$

Donde: Em= error de medida y r_{pi} es la correlación entre los puntajes de los ítems pares e impares, corregidas con la fórmula de Sperman-Brown.

El monto de los errores de medida indica la amplitud del intervalo confidencial en las normas de las edades de vocabulario. Por ejemplo para la edad verbal de 5 años el error de medida es igual a 3, por lo tanto el intervalo confidencial debe ser igual a ± 3 meses (entre 4 años y 9 meses y 5 años y 3 meses). En la siguiente tabla aparece para cada grupo de edad el coeficiente de confiabilidad y el error de medida

Tabla 6: Coeficientes de Confiabilidad y Error Típico de Medida por grupo de edad.

Edades	Coeficiente de Confiabilidad	P	Error Típico de Medida
2.06-3.05	.88	0.000	5
3.06-4.05	.91	0.000	5
4.06-5.05	.93	0.000	4
5.06-6.05	.95	0.000	3
6.06-6.05	.94	0.000	4
7.06-7.05	.90	0.000	5
8.06-9.05	.80	0.000	7
9.06-10.05	.93	0.000	4
10.06-11.05	.94	0.000	4
11.06-12.05	.93	0.000	4
12.06-13.05	.92	0.000	4
13.06-14.05	.93	0.000	4
14.06-15.05	.88	0.000	5
15.06-16.05	.94	0.000	4
16.06-17.06	.60	0.002	9

En todos los grupos de edad la confiabilidad es significativa estadísticamente. El valor de la mediana del coeficiente de confiabilidad es .93 y del error de media 4. Dunn, (1965), en la normación de la prueba

en la población norteamericana, reporta que la mediana del coeficiente de confiabilidad es de .77 y el error de medida 7.2. En la normación realizada en España, Dunn, (1985) obtuvo que la mediana del coeficiente de confiabilidad fue igual a .96 y la mediana del error de medida 3.

En nuestra muestra el coeficiente de confiabilidad más bajo apareció en el último grupo de edad 16.6-17.05 ($r=.60$), pero es significativo estadísticamente. El mayor coeficiente de confiabilidad aparece en el grupo 5.06-6.05 ($r=.95$).

En la otra estimación que se hizo de la confiabilidad de la prueba se utilizó la técnica de test-retest. Para ello se seleccionaron 32 niños que no participaron en el estudio normativo (13 varones) con una edad promedio de 9.19 años ($ds=0.31$). En este estudio se le aplicó la prueba a los niños y pasado un mes se les volvió a aplicar. Se calculó el coeficiente de correlación de Pearson entre el puntaje obtenido en la primera y la segunda aplicación. El valor del coeficiente de confiabilidad obtenido en esta forma fue .72 ($p=0.000$) y el error de medida fue igual a 7.19. El coeficiente de confiabilidad y el error de medida son similares a los obtenidos en el estudio test-retest de la normación española de la PVP ($r=.77$ y error de medida=7).

8. Estimación de la validez

Para determinar la validez de la prueba se hicieron dos pruebas. Primero se correlacionó el puntaje obtenido en la PVP con la cantidad de palabras producidas en el Inventario de Primeras Palabras (IPP) de Manzano, Piñeiro e Inguanzo (1998). El IPP fue diseñado para el estudio del vocabulario de niños de habla hispana de entre 1 y 3½ años de edad.

La primera parte del IPP es una lista de 414 palabras de diferente tipo de gramatical y categoría. En este instrumento los padres deben marcar las palabras que sus hijos producen. Tanto la PVP como el IPP evalúan el vocabulario por lo tanto los resultados que obtenga un niño en una prueba deben estar relacionados con los obtenidos en la otra. Se les aplicó la PVP y el IPP a 26 niños del grupo normativo con edades entre los 3 y 3 años y 5 meses (media de edad=3.02 $ds=0.16$). La correlación obtenida entre el puntaje de la PVP y el IPP fue .44 ($p=.02$). Esta correlación es significativa estadísticamente, lo que indica que ambas pruebas están midiendo un mismo proceso: el vocabulario. Los resultados anteriores son similares a los obtenidos por Bates y colaboradores (1988) cuando correlacionaron la PVP y el CDI Words (inventario de palabras de la lengua inglesa, creado por Fenson y colaboradores, 1994), la correlación obtenida por estos autores fue igual a .51.

El otro estudio de validez se hizo correlacionando los C.I. obtenidos en la PVP con los obtenidos en el WISC-R, (Weschler, 1974). Ambas pruebas se les aplicaron a un grupo de 40 niños (18 varones). Veintiocho niños también participaron en el estudio de test-retest. Se aplicaron todas las tareas del WISC. La correlación entre el C.I. obtenido en la PVP y el C.I. obtenido en la escala verbal del WISC fue .51 ($p=0.001$). La correlación entre el CI de la PVP y el CI de la Escala ejecutiva fue .49 ($p=0.001$). La correlación entre el C.I. de la PVP y el CI total del WISC fue .52 ($p=0.000$). En las escalas la mayor correlación se obtuvo entre el C.I. del PVP y el C.I. de la Escala Verbal. Este resultado es comparable a otros reportados en la literatura, por ejemplo Dunn, (1985) presenta una serie de más de 60 estudios donde se correlacionó los CI de la PVP y el WISC. Para la Escala Verbal las correlaciones oscilaron entre .04 y .88 (mediana .66). Cuando se correlacionó la prueba completa los coeficientes de correlación oscilaron entre -.16 y .91 (mediana .64). Sattler, (1977), también reseña un conjunto de trabajos donde se correlacionaron los CI de la PVP y el WISC. Para la escala verbal los coeficientes de correlación oscilaron entre 0.36 y 0.94 (mediana .66), estas fueron las correlaciones más altas. Las correlaciones con la escala ejecutiva fueron las más bajas, oscilaron entre 0.21 y 0.74 (mediana 0.54). Las correlaciones con el CI total del WISC oscilaron entre 0.30 y 0.84 (mediana 0.63).

CONCLUSIONES

La PVP es una buena prueba para abordar en la práctica el estudio del desarrollo del vocabulario, permite obtener un CI de una forma sencilla y rápida, aunque sin suplir los resultados de otras pruebas de inteligencia. Para poder utilizar eficientemente la PVP se debe contar con normas que contemplen las variaciones lingüísticas y culturales. Es por ello que las normas de la PVP construidas para los sujetos norteamericanos y para sujetos hispanoparlantes de otras culturas no deben ser utilizadas en la población cubana. El presente trabajo trata de suplir las carencias de datos normativos de la PVP que existían para nuestra población. Los estudios de la correlación entre la edad y los resultados de la prueba, los coeficientes de confiabilidad y validez aquí presentados son evidencias sólidas en favor de las normas cubanas. Se está trabajando en la automatización de la PVP para una aplicación más rápida y homogénea y aprovechar las facilidades que brinda la computación.

AGRADECIMIENTOS

Los autores quieren expresar su agradecimiento a Leidys Pérez por su participación en la recogida de los datos del estudio normativo, en el estudio de la confiabilidad y de la validez.

REFERENCIAS

- BATES, E.; I. BRETHERTON y L. SNYDER (1988): *From first words to grammar: individual differences and dissociable mechanisms*. New York: C.U.P.
- DUNN, L. (1959): *Peabody Picture Vocabulary Test: Original Edition*. Circle Pines, MN: American Guidance Service.
- _____ (1965): *Peabody Picture Vocabulary Test*. Circle Pines, MN: American Guidance Service.
- DUNN, L. y L. DUNN (1981): *Peabody Picture Vocabulary Test-Revised*. Circle Pines, MN: American Guidance Service.
- DUNN, L. (1985): *Test de vocabulario en imágenes Peabody adaptación española*. Madrid: Mepsa.
- DUNN, L.; L. DUNN; C. WHETTON and PINTILLED (1982): *British Picture Vocabulary Scale*. Windsor, U.K.: NFER-NELSON.
- DUNN, L.; E. PADILLA; D. LUGO and L. DUNN (1986): *Test de Vocabulario en imágenes Peabody-Adaptación Hispanoamericana [Peabody Picture Vocabulary Test-Latin American adaptation]*. Circle Pines, MN: American Guidance Service.
- FENSON, L.; P. DALE; J. S. RESNICK; E. BATES and D. THAL (1994): "Variability in early communicative development". *Monographs of the Society for Research in Child Development*, 242, 58.
- GUIFORD, J.P. (1950): *Fundamental Statistics in Psychology and Education*. NY: McGraw-Hill.
- MANZANO, M.; A. PIÑEIRO y G. INGUANZO (1998): *Inventario de Primeras Palabras*. Manual de usuario.
- SATTLER, J. (1977): *Evaluación de la inteligencia infantil*. México. Manual Moderno.
- WECHSLER, D. (1974): *Manual of the Wechsler Intelligence Scale for Children-Revised*. NY: Psychological Corp.