

ALGUNAS CARACTERISTICAS Y POSIBILIDADES DEL TEST DE MATRICES PROGRESIVAS DE RAVEN

Dr. Eduardo Cairo Valcárcel¹, Lic. Eric Cairo Martínez, Dr. Carlos Bouza² y T. Ponce Solozabal¹

¹ Profesor Titular, Facultad de Psicología, Universidad de La Habana, Cuba

² Profesor Titular, Facultad de Cibernética-Matemática, Universidad de La Habana, Cuba

RESUMEN

En el trabajo se presentan los resultados de una investigación con 589 sujetos a los cuales se les aplicó el Test de Matrices Progresivas de Raven y se hace una evaluación de algunas de sus características y posibilidades ofreciéndose como alternativas variantes que pudieran ser igualmente eficientes pero a su vez muchos más ahorrativas en el tiempo de aplicación y, por ende, en el costo de su aplicación.

ABSTRACT

In this paper we present an evaluation of some of the characteristics and possibilities of Raven's Progressive Matrices Test and an alternative that can be equally efficient as the original test but would be beneficial in lowering the costs of application due to the economy of time.

INTRODUCCION

El test de Matrices Progresivas de L.S. Penrose y J.C. Raven constituye hoy en día una de las pruebas de inteligencia de mayor empleo en el campo de la Psicometría debido principalmente a varias razones y, entre ellas, a las siguientes:

- Es apropiado tanto para niños mayores de 5 años como para adultos.
- Se ajusta perfectamente no sólo a la edad del sujeto sino y también a sus posibilidades intelectuales y nivel instruccional incluso el mismo puede ser aplicado valiéndonos de la demostración sin el uso del lenguaje si fuera necesario.
- Su empleo puede ser tanto individual como colectivo lo cual ofrece una opción inconmensurable tanto para el trabajo en el área clínica, generalmente individual, como para el trabajo en el campo de la Psicología Educacional y Laboral donde con frecuencia se requiere su uso en forma colectiva.
- Su bajo componente verbal lo cual implica en la práctica una ventaja grande cuando se requiere su utilización en poblaciones con bajo o muy bajo nivel instruccional, aunque hay investigaciones donde se señala la necesidad de no considerar esta propiedad de la prueba (bajo nivel de exigencia de comprensión verbal) como un valor absoluto.
- Se considera como una medida confiable del factor "g" de Spearman.

En nuestro país su utilización está muy difundida aunque el aprovechamiento de sus múltiples alternativas no siempre son empleadas a fondo, al menos, queremos decir, en el caso de que se utilicen sus resultados no han sido difundidos lo necesario.

Por tales razones, y debido al empleo continuo de esta prueba, hemos deseado compartir una parte de nuestra experiencia y algunas reflexiones vinculadas con su uso y posibilidades. Asimismo, mostrar como hemos enfrentado varias de las críticas que con más frecuencia se le hace a la prueba. Ellas son: No contar con normas lo suficientemente adecuadas así como datos mínimos sobre su confiabilidad y validez (Kaplan and Sacuzzo 1988) aunque esto no significa, en modo alguno, se considere satisfecha con este trabajo, la necesidad presente de hacer una verdadera y representativa estandarización en una muestra cubana.

Así, el trabajo tiene los siguientes objetivos:

- Establecer el índice de dificultad y discriminación de cada ítem de la prueba.
- Mostrar algunas alternativas del empleo de las dos formas básicas de calificación (Por los aciertos y por los errores).
- Proponer nuevas alternativas de presentación de la prueba a partir de nuestra experiencia.
- Presentar las normas de percentil obtenidas en nuestro trabajo.

¹Email cairo@psico.u.h.cu

Metodología

I- Características de la muestra.

La muestra estudiada está compuesta por 589 sujetos de ellos 286 hombres y 303 mujeres con una edad promedio de 25,869 años en un rango de 57 que va de 11 a 68 años, con una desviación estándar de 8,965, un error estándar de la media de .369 y una escolaridad media de secundaria básica.

En la Tabla # 1 pueden apreciarse diferentes características de la muestra investigada y entre ellas las siguientes:

- La edad media de los grupos escolares tiende a aumentar a medida que se incrementa el nivel instruccional aunque el contingente de preuniversitario presenta una media de edad mayor que el de los universitarios (26.1 y 24.4 respectivamente).
- Vale aclarar que estos dos grupos referidos son los que muestran una desviación estándar menor (8.1) lo que significa en la práctica que en ambos grupos los sujetos se encuentran más agrupados alrededor de la media que en los restantes.

- La edad media de la muestra general es de 25,869 años como ya fue señalado y la diferencia media de edad entre los hombres y las mujeres es mínima (0.3 años). Sin embargo, el grupo de los hombres presenta una variabilidad (desviación estándar) mucho menor que el de las mujeres (8.4 los hombres y 9.4 las mujeres) lo que significa que el primero es más homogéneo en relación con este parámetro.

II- Técnica utilizada.

Test de Matrices Progresivas de Raven forma general o estándar el cual permite medir habilidades del pensamiento abstracto, habilidad para resolver nuevos problemas, capacidad de observación, razonamiento y analogías.

III- Procedimiento:

- a- La aplicación de la prueba fue llevada a cabo por personal profesional y/o estudiantes universitarios entrenados para tal objetivo para garantizar la homogeneidad del proceder.

Tabla 1. Distribución de la muestra según la escolaridad, la edad y el sexo.

Grupo Escolar	Grupos Etáreos					Total	Media	DS	ESM
	<21	21-30	31-40	41-50	>50				
Second.	68	62	15	2	5	152	23.8	9.1	0.739
Hom.	33	31	6	1	2	73	23.7	9.0	1.056
Muj.	35	31	9	1	3	79	23.9	9.2	1.041
Preuniv.	49	118	37	6	5	215	26.1	8.1	0.550
Hom.	25	60	19	2	1	107	25.4	6.8	0.660
Muj.	24	58	18	4	4	108	26.8	9.1	0.875
Univers.	47	94	10	4	5	160	24.4	8.1	0.643
Hom.	14	49	3	2	3	71	24.9	8.3	0.985
Muj.	33	45	7	2	2	89	24.1	8.0	0.852
Profes.	0	30	16	14	2	62	33.9	9.3	1.182
Hom.	0	18	10	5	2	35	32.7	8.8	1.484
Muj.	0	12	6	9	0	27	35.5	9.9	1.903
Total	164	304	78	26	17	589	25.9	9.0	0.369
Hom.	72	158	38	10	8	286	25.7	8.4	0.499
Muj.	92	146	40	16	9	303	26.0	9.4	0.543
Edad Media	18.3	24.3	34.6	45.7	57.5	-	-	-	-
Hom.	18.1	24.3	34.4	44.9	56.5	-	-	-	-
Muj.	18.4	24.2	34.7	46.1	58.4	-	-	-	-
DS	1.8	2.6	2.9	2.8	5.6	-	-	-	-
Hom.	1.8	2.5	2.7	2.6	4.4	-	-	-	-
Muj.	1.8	2.7	3.0	3.0	6.7	-	-	-	-
ESM	0.14	0.15	0.33	0.56	1.37	-	-	-	-
Hom.	0.22	0.20	0.44	0.82	1.56	-	-	-	-
Muj.	0.19	0.22	0.48	0.74	2.23	-	-	-	-

Secund=Secundaria Básica,
Prof=Profesionales,

Preuniv=Preuniversitario,
Hom=Hombres, Muj=Mujeres

Univers=Universidad

- b- La prueba se aplicó de forma individual o colectiva según fuera la necesidad de la situación.
- c- Todos los sujetos fueron instruidos acorde con las orientaciones establecidas por los autores de la prueba. (Raven, 1970)
- d- La calificación de los tests se efectuó manualmente y se verificó el resultado de cada evaluación mediante un programa computacional creado al efecto tanto para la valoración de los aciertos como de los errores.

Asimismo, se consideró por los autores que el orden de los ítems dentro de cada serie estaba también ordenado en relación con su dificultad de tal forma que el ítem # 1 de cada serie fuera más fácil que el ítem # 2, y así sucesivamente hasta el ítem # 12.

Ahora bien, no obstante esta suposición teórica, nuestros resultados no se ajustan exactamente a la misma pues como puede observarse en la Tabla # 2 que más abajo insertamos no siempre ocurrió lo previsto por el autor.

Tabla 2. Orden de dificultad de cada ítem de la prueba sin y con corrección del azar. (En por ciento).

Item #	Serie s/c	A c/c	Serie s/c	B c/c	Serie s/c	C c/c	Serie s/c	D c/c	Serie s/c	E c/c
1	100.0	100.0	99.7	99.6	97.1	96.7	97.3	96.9	80.6	77.9
2	99.5	99.4	99.2	99.0	93.2	92.2	93.4	92.4	66.2	61.4
3	99.5	99.4	98.1	97.8	91.5	90.3	92.4	91.3	69.3	64.9
4	99.7	99.6	94.9	93.9	86.1	84.1	87.6	85.8	54.7	48.2
5	99.0	98.8	94.4	93.3	89.1	87.6	94.6	93.8	57.7	51.7
6	99.3	99.2	89.0	86.8	72.0	68.0	89.6	88.2	52.5	45.7
7	94.9	93.9	83.0	79.6	87.8	86.0	83.9	82.1	37.2	28.2
8	95.1	94.1	74.7	69.6	65.4	60.4	83.0	80.6	33.4	23.9
9	98.1	97.8	75.7	70.9	72.7	68.8	77.2	76.4	29.7	19.7
10	94.1	92.9	84.6	81.5	60.6	55.0	73.7	69.9	18.8	7.3
11	85.4	82.5	76.9	72.3	46.9	39.3	40.6	32.1	12.9	0.5
12	76.2	71.5	57.0	48.5	20.9	9.6	18.2	6.5	12.6	0.1
Pro	95.1	94.1	85.6	82.7	73.6	69.8	77.6	74.7	43.8	35.8
Diferencia (s/c-c/c)		1.0	2.9	3.8	2.9	8.0				

s/c= Sin corrección c/c= Con corrección

- e- Finalmente, se efectuaron todos los análisis estadísticos mediante el SPSS/PC Versión 8.0 y se procedió a los estudios que más abajo detallamos.

Análisis de los resultados

I- Análisis de ítems

El primer paso que se efectuó fue determinar el índice de dificultad para cada ítem lo cual constituyó una simple operación aritmética debido al hecho de haberle facilitado a los sujetos sometidos a estudios el tiempo necesario para poder considerar todos los ítems lo que permitió poder considerar que el puntaje más alto posible pudiera obtenerse disminuyendo del total de respuestas la cantidad de soluciones erróneas y preguntas dejadas en blanco.

A -Determinación del índice de dificultad o comprensibilidad de cada ítem.

Como se conoce el test de Matrices Progresivas para Adultos de J.C.Raven está concebido de tal manera que los 60 ítems de que consta están agrupados en cinco series diferentes de dificultad creciente que son: A, B, C, D y E.

En la Tabla # 2 antes referida pueden destacarse varios aspectos y entre ellos los siguientes:

El ítem más fácil es el A1 y el más difícil es el E12 con un 100 % y un 12.6 % de respuestas correctas respectivamente.

Asimismo, se destaca como la serie A resulta la más fácil con un 95.1 % y la serie E la más difícil con un 43.8 % todo lo cual nos indica que los parámetros más generales y extremos concebidos por los creadores de la prueba se conservan. No obstante, debe indicarse al igual que fue subrayado en los trabajos de Cairo (1997, 1998) y Ponce Solozabal (1999) que la estructura interna de cada una de las series así como el orden de las series debe ser motivo de revisión por cuanto tanto una como otra muestran alteraciones que no se corresponden con la organización inicial considerada para la estructura de la prueba. Varios ejemplos de ellos son:

En la serie A el ítem 7 (94.9 % de aciertos) resulta más difícil que el ítem 8 (95.1 % de aciertos) y el ítem 9 (98.1 % de aciertos) resulta más fácil que el 7 y el 8. Casos semejantes se dan dentro de las series B, C, D y E.

Otro aspecto importante que debe ser tomado en cuenta es la revisión y ordenación de las series C y D por cuanto en varias investigaciones como fue apuntado con anterioridad, resulta que el promedio de aciertos en la serie D (77.6 %) resulta más alto que en la serie C (73.6 %).

En la Tabla # 2 también podemos observar que el índice de dificultad pudiera clasificarse en tres categorías o niveles de acuerdo con Garret (1974). Así, llamaríamos ítems fáciles a los que son resueltos por más del 60 % de los sujetos y entonces resultarían dentro de esta categoría 46 ítems (76.7 %), denominaríamos ítems moderados a aquellos que son resueltos favorablemente entre el 40 y el 60 % y entonces tendríamos en este nivel 6 ítems (10 %) y, finalmente, en la categoría de ítems difíciles se ubicarían 8 ítems (13.3 %).

Del resultado anterior concluimos que la prueba en general resulta fácil para la población en que fue utilizada lo cual nos confirma la necesidad de mejorar su estructura, de tal forma, que permita evaluar más diferencialmente a los sujetos pertenecientes a poblaciones con características semejantes.

Otro aspecto que puede apreciarse en la Tabla # 2 es que la aplicación de la fórmula para determinar el índice de dificultad con corrección al azar (Lienert, 1990) no altera significativamente, en este caso, los resultados hallados. No obstante, vale señalar que a medida que se dificulta la tarea al aumentar el nivel de complejidad de los ítems en cada serie se estimula, en cierta medida, la posibilidad de emplear el azar en la búsqueda de la solución requerida.

Así, de acuerdo con lo anterior vemos que las diferencias obtenidas en las distintas series aumenta de 1.0 en la serie A hasta 8.0 en la serie E aunque es válido subrayar que las alternativas de respuestas posibles en las dos primeras series (A y B) es de 6 mientras que en las otras series (C, D y E) es de 8.

En la Tabla # 3 se hace un resumen de la Tabla # 2 y se aprecia claramente que las diferencias al clasificar los ítems mediante la utilización o no de la fórmula de la corrección por el azar no resultó en la muestra estudiada significativa.

Tabla 3. Clasificación de los ítems según el índice de dificultad en la muestra total.

Series	<40		Entre 40 y 60		>60		Total	
	s/c	c/c	s/c	c/c	s/c	c/c	s/c	c/c
A	-	-	-	-	12	12	12	12
B	-	-	1	1	11	11	12	12
C	1	2	1	1	10	9	12	12
D	1	2	1	0	10	10	12	12
E	6	6	3	3	3	3	12	12
Total	8	10	6	5	46	45	60	60

s/c=Sin corrección c/c=Con corrección

Como posible solución de estos problemas estamos investigando en tres direcciones factibles, al menos, en estos momentos.

La primera consiste en la eliminación de la organización serial mediante una propuesta que presenta de corrido todos los ítems ordenados jerárquicamente, en orden creciente de dificultad, y reduce la prueba en el número de ítems a 45.

Lo anterior permite, a nuestro juicio, contribuir a resolver algo que en estos momentos coyunturales de nuestro desarrollo económico se plantea como una necesidad del Perfeccionamiento Empresarial en la Empresa Estatal Cubana referido entre otros aspectos a:

- Evaluar la eficiencia en cuanto al uso de los recursos materiales, financieros y de la fuerza de trabajo, que se emplean en la actividad.
- Facilitar la valoración de posibles decisiones a tomar, que permitan la selección de aquella variante, que brinde el mayor beneficio con el mínimo de gastos.
- Analizar la posibilidad de reducción de gastos.

-y también expresado como recomendación en la tesis de Ponce Solozabal (1999, Pág. 94) cuando refiere: "Consideramos oportuno, dado el tiempo que se invierte en el proceso de selección de los candidatos, estudiar particularmente a profundidad el tiempo que se emplea en la ejecución de esta prueba y la posibilidad de simplificarla conservando su capacidad diagnóstica y, a la vez reduciendo el gasto económico que significa su aplicación".

Finalmente, creemos que debemos insistir hasta la saciedad en la imperiosa necesidad de no sólo conservar y mejorar el nivel profesional con que aplicamos las pruebas y desarrollamos nuestro trabajo sino y, también, no olvidar que los gastos en que incurrimos en la prestación del servicio constituye una muestra de la eficiencia con que se logran los resultados. Así, contribuimos como señalábamos más arriba con la aplicación de uno de los principios generales del Perfeccionamiento

Empresarial que plantea: "El perfeccionamiento empresarial es un proceso de mejora continua de la gestión interna de la empresa que posibilita lograr de forma sistemática, un alto desempeño para producir bienes o prestar servicios competitivos". (Comité Ejecutivo del Consejo de Ministros, s/f, pág. 3).

Las normas percentiles para esta variante serían las siguientes:

Percentil	Calificación	Percentil	Calificación	Percentil	Calificación
95	41	60	32	25	26
90	39	55	32	20	24
85	37	50	31	15	23
80	36	45	30	10	20
75	35	40	29	5	15
70	34	35	28		
65	33	30	27		

También vale subrayar que con esta proposición el índice de dificultad medio disminuyó de 75.1 % (Ver Tabla # 2) a 71.1 % y el índice de discriminación medio aumentó de .3846 a .4306 aspectos ambos beneficiosos para la efectividad de la prueba.

La **segunda alternativa** en la cual estamos laborando consiste en la reestructuración completa de las series C y D lo cual ha implicado no sólo reordenar los items sino también pasar algunos de una serie a la otra.

A partir de los resultados de este trabajo y de los otros antes efectuados estamos investigando con la siguiente nueva proposición de organización de la serie C y D.

Nueva proposición de la serie C:

Item 1 D1 Item 2 D5 Item 3 C2

Item 7 D7 Item 8 D9 Item 9 C9

Item 4 C3 Item 5 C5 Item 6 D4

Item 10 C8 Item 11 C11 Item 12 C12

Nueva propuesta de la serie D:

Item 1 C1 Item 2 D2 Item 3 D3

Item 7 D8 Item 8 D10 Item 9 C6

Item 4 D6 Item 5 C7 Item 6 C4

Item 10 C10 Item 11 D11 Item 12 D12

Organizados así las series C y D tendrían un índice de dificultad promedio de 76.7 y 74.5 respectivamente.

Su índice de discriminación medio resultaría entonces de .4539 y .4569 para cada una de estas series.

Con esta nueva variante se lograría, a nuestro entender, ajustar la prueba a la concepción de dificultad creciente concebida por los autores y, a la vez, algo que desde el punto de vista psicológico repercutiría favorablemente en el ánimo del sujeto y es eliminarle la tensión de la inestabilidad que produce el someterse a una situación experimental donde el algoritmo organizativo no le permite establecer una expectativa clara, evidente, deducible en cuanto a las características de complejidad del item que debe resolver.

En otras palabras, con esta organización serial el sujeto sabe, pues deduce fácilmente que cada nuevo item presentará una dificultad mayor que el anterior.

La **tercera alternativa** sobre la cual trabajamos consiste en la sustitución, reordenamiento y completamiento de un grupo más o menos numeroso de items lo cual nos permite en una sola prueba incluir tanto los requerimientos exigidos para una población adulta de nivel instruccional bajo o medio como para

profesionales de nivel superior eliminando a su vez la organización serial.

Esta última variante trae aparejada algo de lo que han adolecido las pruebas de Matrices Progresivas y es la de contar con normas lo suficientemente amplias y representativas. Sin embargo, dado que nos encontramos en la parte inicial del pilotaje y/o experimentación de las ideas básicas creemos que esta dificultad podrá superarse.

B- Determinación del índice de discriminación o validez interna de cada item

La determinación del índice de discriminación o de homogeneidad constituye un paso fundamental en todo proceso de análisis de item de una prueba y este, puede efectuarse mediante el cálculo de diferentes estadísticos o índices.

Uno de ellos, como se conoce, se define a través del coeficiente de correlación r biserial o punto biserial el cual señala el vínculo existente entre las respuestas dadas por el sujeto a un item particular (i) y la puntuación total X del test de la que el item (i) forma parte como elemento.

Vale aclarar, que en general, la correlación biserial puntual se establece como una correlación de Pearson entre dos variables, con la particularidad de que una de esas variables es de tipo continuo y la otra es una variable dicotómica. (Santiesteban Requena, 1990)

Nosotros utilizamos la fórmula propuesta por el autor antes referido (Santiesteban Requena 1990, pág. 340) y conseguimos los resultados más abajo expuestos en las Tablas 4 y 5.

Según puede apreciarse en los resultados de la Tabla # 4 solamente hay tres items que no discriminan (A1, A5 y B1) y dos items que discriminan exclusivamente al 0.05 %, ellos son: el A2 y el A3.

Estos datos son semejantes a los obtenidos por Ponce Solozabal quien reportó 10 items no significativos donde se encuentran incluidos los reportados por nosotros (A1, A5 y B1) y los dos que nosotros informamos que discriminan sólo al 0.05 % (A2 y A3).

El resto de los índices de discriminación determinados resultaron significativos al 0.01 por ciento.

Tabla 4. Índice de discriminación en cada una de las series de la prueba

Item	Serie A	Serie B	Serie C	Serie D	Serie E
1	----(n)	.0703(n)	.3629	.3649	.5339
2	.1005(x)	.2199	.4157	.4605	.5524
3	.0862(x)	.1902	.4501	.4268	.5493
4	.1297	.3586	.5111	.4909	.5625
5	.0475(n)	.3889	.5369	.4957	.5790
6	.1194	.4335	.5136	.5209	.5067
7	.3327	.3796	.4836	.4667	.3436
8	.2611	.4655	.5045	.4769	.4211
9	.2457	.5111	.4495	.4778	.4502
10	.3956	.4837	.3804	.5377	.3352
11	.2609	.5303	.3951	.4309	.2445
12	.3001	.5169	.3989	.3770	.2382
Prom	.1900	.3790	.4502	.4606	.4431

N=No significativa X=Significativa al 0.05

II. Índice de consistencia de la puntuación final.

No podríamos dejar de mencionar que los datos sobre los cuales hemos estado trabajando cumplen los requisitos señalados por Raven (1970) en cuanto al índice de consistencia de la puntuación final de cada sujeto lo cual mostramos en las distintas agrupaciones que hemos efectuado para los análisis respectivos.

Como podrá observarse en la tabla anterior el índice de consistencia de la puntuación final en cada uno de los grupos estudiados y en la muestra total se mantiene dentro de los parámetros

Tabla 5. Resumen del índice de discriminación en cada serie

Índice Discriminación	Serie A Item No.	Serie B Item No.	Serie C Item No.	Serie D Item No.	Serie E Item No.	Totales	
						Cantidad	Por ciento
0 y 10	4	1	-	-	-	5	8.3
11 y 20	2	1	-	-	-	3	5.0
21 y 30	4	1	-	-	2	7	11.7
31 y 40	2	3	4	2	2	13	21.7
41 y 50	-	3	5	8	2	18	30.0
51 y 60	-	3	3	2	6	14	23.3
> 60	-	-	-	-	-	-	-
Totales	12	12	12	12	12	60	100.0

Además, queremos destacar que el cálculo del índice de discriminación pudo haberse efectuado mediante otros procedimientos de cómputo y entre ellos a través de las tablas de J.C. Flanagan práctica muy acostumbrada pues se juzga muy conveniente para ver el grado hasta donde son capaces de dar resultados semejantes fórmulas y quehaceres diferentes. (Garrett 1974)

También consideramos oportuno subrayar que el índice de discriminación está vinculado con el índice de dificultad pues este varía no sólo a partir de la obtención de altos por cientos de aciertos (Ejemplo: Un item muy fácil todos los contestan y, por ende, no diferencia entre los más aptos y los menos aptos) sino y también cuando se obtienen muy bajos por cientos de aciertos (Ejemplo: Cuando tanto los menos aptos como los más aptos no logran contestar correctamente).

Por eso se recomienda como items óptimos aquellos que representan un nivel aceptable tanto de su capacidad discriminativa como de su asequibilidad para los examinados (Kaplan and Sacuzzo, 1985), aspecto este que tuvimos en cuenta en la proposición de reducción de la prueba 1 45 items (Ver primera alternativa).

considerados por Raven (1970) como normales lo cual permite, por ende, tener la confianza necesaria para estimar que la capacidad general determinada por la prueba en cada uno de los grupos es representativa de determinado nivel intelectual que le permite a los sujetos desarrollar una actividad inteligente.

III- Evaluación de la prueba acorde con la cantidad y magnitud de los errores cometidos por los sujetos

En 1950 el Departamento de Psicología Educacional de la Plata, Argentina publicó la alternativa de evaluar los resultados de esta prueba mediante la calificación de la magnitud de los errores idea que sustentaron sobre la necesidad que tiene el evaluado de elegir entre 6 u 8 opciones posibles una respuesta que sea adecuada lo que trae consigo con frecuencia que se seleccionen alternativas erróneas por no contemplar estas respuestas todos los requisitos indispensables.

Algunas respuestas erróneas tienen algunos elementos (1 o 2) pero no los tres necesarios para conseguir la respuesta correcta, de tal forma, que la proposición de evaluación de los errores consiste en síntesis en considerar los errores en una escala de tres dimensiones las cuales son: Errores Finos se

Tabla 6. Evaluación promedio obtenida en la investigación y calificación esperada según Raven (1970)

	Serie A	Serie B	Serie C	Serie D	Serie E	Total
I. En la muestra total						
Evaluación Promedio obtenida	11	10	9	9	5	44
Calificación esperada según Raven	12	10	9	9	4	44
Diferencia	+1	0	0	0	-1	0
II. En la muestra de mujeres						
Evaluación Promedio obtenida	11	10	9	9	5	44
Calificación esperada según Raven	12	10	9	9	4	44
Diferencia	+1	0	0	0	-1	0
III. En la muestra de hombres evaluación						
Promedio obtenida Calificación	11	10	9	9	5	44
esperada según Raven Diferencia	12	10	9	9	4	44
	+1	0	0	0	-1	0
IV. En la muestra de secundaria básica						
Evaluación Promedio obtenida	11	9	8	8	4	40
Calificación esperada según Raven	11	10	8	8	3	40
Diferencia	0	+1	0	0	-1	0
V. En la muestra de preuniversitario						
Evaluación Promedio obtenida	11	10	9	9	5	44
Calificación esperada según Raven	12	10	9	9	4	44
Diferencia	+1	0	0	0	-1	0
VI. En la muestra de universitario						
Evaluación Promedio obtenida	12	11	10	10	7	50
Calificación esperada según Raven	12	11	10	10	7	50
Diferencia	0	0	0	0	0	0
VII. En la muestra de profesionales						
Evaluación Promedio obtenida	12	11	10	10	7	50
Calificación esperada según Raven	12	11	10	10	7	50
Diferencia	0	0	0	0	0	0
VIII. En la muestra de menores de 21 años						
Evaluación Promedio obtenida	11	10	9	9	5	44
Calificación esperada según Raven	12	10	9	9	4	44
Diferencia	+1	0	0	0	-1	0
IX. En la muestra de 21 a 30 años						
Evaluación Promedio obtenida	11	10	9	9	5	44
Calificación esperada según Raven	12	10	9	9	4	44
Diferencia	+1	0	0	0	-1	0
X. En la muestra de 31 a 40 años						
Evaluación Promedio obtenida	11	10	8	9	4	42
Calificación esperada según Raven	11	10	9	9	3	42
Diferencia	0	0	+1	0	-1	0
XI. En la muestra de 41 a 50 años						
Evaluación Promedio obtenida	11	10	9	9	6	45
Calificación esperada según Raven	12	10	9	9	5	45
Diferencia	+1	0	0	0	-1	0
XII. En la muestra de mayores de 50 años						
Evaluación Promedio obtenida	11	10	8	9	5	43
Calificación esperada según Raven	12	10	9	9	3	43
Diferencia	+1	0	+1	0	-2	0

consideran aquellos que sólo tienen en cuenta dos de los tres elementos requeridos, Errores Medios los que consideran la educación de un solo error y Errores Gruesos los que no tienen en cuenta la educación de ningún elemento.

Considerado así el procedimiento, hicimos los cálculos pertinentes y alcanzamos los resultados que se presentan en la tabla # 7 incluida más abajo.

En la tabla antes referida puede apreciarse como el por ciento de error medio alcanza la suma de

Tabla 7. Resumen de los tipos de errores cometidos por los sujetos en las cinco series. (En por ciento)

	Serie A	Serie B	Serie C	Serie D	Serie E	Promedio
Grueso	1.51	3.21	5.63	5.79	27.83	8.80
Medio	1.67	3.31	8.04	5.97	7.77	5.35
Fino	1.73	4.26	11.95	9.14	8.45	7.11
No clas	-	3.42	-	-	0.85	0.85
Total	4.91	14.20	25.62	20.90	44.90	22.11
Omis.	0.03	0.20	0.78	1.49	11.30	2.76
Aciertos	95.06	85.60	73.60	77.61	43.80	75.13
Total General	100.00	100.00	100.00	100.00	100.00	100.00

22.11 el cual se expresa en concordancia con el nivel de dificultad de cada una de las series excepto la C y la D las cuales no cumplen con la idea inicialmente concebida de dificultad creciente concebida por los autores de la prueba y que fue ya señalado con anterioridad en este trabajo y en otros previamente dados a conocer. (Cairo 1998, Ponce 1999)

Asimismo, vale la pena señalar como en la medida en que aumenta el nivel de complejidad de la serie aumentaron los errores y, especialmente, los gruesos quienes crecen desde 1.51 en la serie A hasta 27.83 por ciento en la serie E.

Estrechamente vinculado con el análisis de los errores se debe destacar que el número de alternativas erróneas empleadas por los sujetos en la búsqueda de la respuesta correcta de cada ítem nos ofrece un indicador no sólo del nivel de dificultad del ítem, sino también del valor de los distractores (5 para cada uno de los ítems de las series A y B) y (7 para los ítems de las series C, D y E).

Ahora bien, en esta ocasión sólo queremos referir que los sujetos por nosotros estudiados tuvieron el siguiente comportamiento:

En la serie A utilizaron el 51.67 % del número de las alternativas erróneas posibles, en la serie B el 68.33 %, en la C el 94.05 %, en la serie D el 92.86 % y en la serie E el 100.00 %.

Finalmente, agregaríamos al respecto que se hace indispensable en la revisión del test incluir el estudio del valor de cada uno de los distractores pues existen muchos de ellos que prácticamente no son seleccionados por los sujetos lo cual indica que no desempeñan ningún rol como tales. No obstante, este aspecto debe ser motivo de un análisis independiente y especial en otra ocasión.

Estos datos obtenidos mediante la evaluación de los errores de cada sujeto nos permiten una alternativa adicional pues nos da la posibilidad de diferenciar entre dos sujetos que obtienen por la vía de la calificación de los aciertos igual evaluación pero que pudieran no haber cometido igual número de errores en cuanto a su magnitud. Esta situación

cobra una mayor importancia en el trabajo de selección de grandes contingentes de personal donde con frecuencia hay muchos sujetos que coinciden con una misma evaluación.

Asimismo, consideramos que el análisis del número de educciones planteadas en cada uno de los ítems donde la respuesta ofrecida por el sujeto ha sido errada no debe ni con mucho concluir con un análisis cuantitativo y/o simplemente estadístico como el referido más arriba sino por el contrario debe constituir el primer eslabón necesario para el vínculo imprescindible de promover el desarrollo, la capacitación y/o entrenamiento del personal.

En este sentido creemos pues nos ha resultado muy útil especialmente en el campo de la promoción del desarrollo a través de la búsqueda de alternativas de respuestas para la solución de problemas la utilización del entrenamiento de los sujetos en el análisis de los elementos como base de ayuda para la solución de determinadas tareas.

Ahora bien, pudiéramos preguntar si obtenida ya la calificación total de la prueba acorde con el número de aciertos y también la calificación total según el número de errores qué opción pudiéramos asumir preferiblemente. Naturalmente, que las respuestas como en infinidad de situaciones en las cuales el psicólogo debe decidir sobre una situación semejante depende en última instancia del objetivo propuesto.

Entonces, debemos reflexionar entre las otras muchas cuestiones sobre dos que tal vez sean interesantes como punto de partida. Ellas son: ¿Cuál sería la mejor elección factible si debemos elegir, de un contingente determinado de sujetos, los más aptos? ¿Serían aquellos que más puntuaciones de acierto obtuvieran mostrando así su capacidad para solucionar tareas que requieren diferentes niveles de complejidad? o ¿Sería preferible elegir aquellos que cometieron menos errores cuando respondieron la prueba?

Estas dos alternativas nos la ofrecen el sistema de calificación por aciertos o errores del Raven. Sin embargo, es preferible en la mayoría de los casos

lograr una combinación donde no sólo se elijan los que mayores puntuaciones de aciertos tienen sino y también cuando la magnitud de sus errores es mínima.

Veamos ahora, un ejemplo tomado de esta muestra el cual pudiera ser representativo para muchos otros trabajos. De hecho, esta idea ya se ha aplicado en la práctica y, en parte, fue comentada en el trabajo de Ponce Solozabal (1999).

Ejemplo No. 1

Pensemos que necesitamos seleccionar de este mismo grupo de 589 sujetos a 159 para cubrir determinados puestos de trabajo los cuales exigen para su óptimo desenvolvimiento la capacidad de solucionar problemas que se presentan continuamente aunque no en déficit de tiempo y, además, los errores y/o decisiones inadecuadas, inoportunas que adopte el trabajador en un momento determinado sólo conllevarán atrasos en la culminación del trabajo no así pérdidas materiales salvo lógicamente las consecuencias económicas que trae consigo la demora.

Podríamos en este caso entonces tomar la decisión de elegir a aquellos que obtuvieran las mejores calificaciones en el Raven y cumplieran lógicamente los otros parámetros exigidos por el puesto de trabajo.

La decisión anterior implicaría que se incorporarán al centro laboral los siguientes 159 nuevos trabajadores:
Trabajador No.

4, 5, 13, 15, 16, 20, 24, 25, 27, 30, 33, 34, 36, 37, 40, 42, 45, 47, 48, 49, 50, 51, 52, 54, 55, 56, 61, 62, 64, 68, 70, 71, 72, 75, 82, 86, 97, 98, 99, 102, 105, 107, 108, 109, 113, 114, 117, 118, 132, 133, 138, 140, 145, 146, 148, 152, 153, 154, 155, 157, 160, 162, 164, 165, 167, 170, 172, 173, 179, 186, 189, 193, 195, 196, 197, 201, 203, 204, 206, 207, 209, 212, 213, 214, 218, 219, 222, 226, 227, 231, 235, 236, 239, 241, 251, 257, 259, 270, 271, 275, 276, 282, 283, 284, 285, 286, 287, 288, 314, 328, 335, 336, 341, 348, 367, 379, 407, 438, 446, 448, 452, 457, 469, 476, 483, 487, 491, 495, 497, 502, 506, 517, 518, 519, 521, 522, 525, 527, 530, 532, 536, 538, 540, 541, 542, 545, 546, 551, 555, 558, 559, 561, 574, 575, 578, 583, 585, 586 y 588.

Estos trabajadores como grupo tendrían una calificación total media de aciertos de 53.925 y una calificación media de errores de 5.135 (desviación estandar de 2.61 y 2.81 y un error estandar de la media de 0.21 y 0.22 respectivamente.

Ejemplo No. 2

Si en lugar de tener la necesidad de seleccionar los aspirantes para cubrir plazas de trabajo como las

anteriores referidas sino elegir operarios elaboradores de determinados productos donde no sólo es necesario tener la capacidad para resolver los problemas que continuamente se presenten, sino es imprescindible también minimizar los errores puestos que los mismos no sólo traen consigo pérdidas económicas por el consiguiente aumento del tiempo de producción sino y también pérdidas materiales muy significativas pues los materiales empleados en la operación fabril del producto son extremadamente costosos y si no se logra con éxito hacer el mismo desde el primer intento hay que desechar totalmente la materia prima empleada ya que esta se deteriora y tomar nuevamente otras porciones para reintentar lograr el objetivo, nosotros nos preguntamos entonces: ¿No sería preferible seleccionar para este caso concreto a partir del criterio de elegir a los aspirantes entre los que comenten menos errores en el Raven y no entre los que obtienen mayores calificaciones de aciertos?

Sin dudas, que existe una correlación alta y negativa entre unos y otros pero no necesariamente resultarían las mismas personas lo cual es un aspecto que pudiera tener consecuencias (favorables o desfavorables) para la empresa si se optara por una decisión u otra.

En nuestro ejemplo # 2 los nuevos 159 trabajadores serían los siguientes:

4, 5, 11, 15, 18, 20, 24, 25, 27, 30, 33, 34, 36, 38, 40, 42, 45, 47, 48, 49, 50, 52, 54, 55, 55, 63, 64, 68, 70, 71, 72, 75, 82, 85, 86, 97, 98, 99, 102, 107, 109, 113, 114, 117, 118, 132, 133, 134, 136, 138, 145, 146, 148, 149, 152, 153, 154, 155, 156, 157, 159, 160, 162, 164, 165, 167, 170, 172, 173, 179, 183, 186, 189, 193, 195, 196, 197, 203, 204, 206, 207, 209, 213, 214, 219, 222, 226, 227, 231, 235, 236, 239, 241, 251, 259, 270, 271, 275, 276, 282, 283, 284, 285, 286, 287, 288, 293, 298, 321, 333, 353, 357, 367, 377, 379, 393, 407, 420, 422, 423, 424, 427, 429, 448, 453, 455, 458, 464, 471, 472, 474, 475, 476, 483, 487, 491, 496, 502, 506, 511, 517, 520, 522, 525, 530, 532, 537, 538, 545, 546, 551, 553, 559, 561, 567, 572, 573, 574, 578, 583 y 585.

Estos trabajadores como grupo tendrían una calificación total media de aciertos de 51.799 y una calificación media de errores de 4.013 (desviación estándar 6.07 y 1.94 y un error estándar de la media de 0.48 y 0.15 respectivamente.

Ahora bien, si comparamos ambas selecciones veremos que no coinciden 41 sujetos, es decir, el 25.8 % de los 159 elegidos.

Por ende, si como psicólogo sólo empleáramos la variante de tener en cuenta exclusivamente la calificación de los aciertos para dar entrada a los nuevos empleados tendríamos éxito en unos casos

y en otros no tanto y viceversa ocurrirá si eligiéramos a los de nueva incorporación por la evaluación de los errores exclusivamente.

De lo anterior se desprende, por tanto, la necesidad de utilizar uno u otro modo de calificación acorde con la necesidad que tengamos y cuando no haya una exigencia específica es preferible elegir al nuevo trabajador a partir de considerar a ambos procedimientos lo cual daría como resultado que el contingente de los 118 sujetos que coinciden en ambas selecciones tendrían una calificación media de aciertos de 54.78 y una calificación media de errores de 3.87 con una desviación estándar de 2.45 y 2.02 y un error estándar de la media de 0.23 y 0.19.

semejanza en cuanto a diferentes atributos, es preferible construir estas normas y decidir mediante ellas. Esa es la razón principal por la cual le incluimos las más abajo referidas.

Aquí debemos señalar que trabajos semejantes han sido efectuados en diferentes instituciones del país antes y después del Triunfo de la revolución y entre ellas se encuentran por ejemplo: Agencia Publicitaria Siboney (1954), CMQ (1955), Facultad de Psicología (Década de 1960), Ministerio de la Pesca (1981) las cuales determinaron sus normas percentiles para las poblaciones con las que trabajaban. (Campos Puentes, s/f)

Asimismo, con el objetivo de satisfacer una

Tabla 8. Tabla de percentiles calculadas sobre los 589 sujetos

Perc.	Muestra Total N=589	Hombres N=286	Mujeres N=203	Secundaria N=152	Preuniversitario N=215	Universitario N=160	Profesionales N=62	< 21 N=164	21-30 N=304	31-40 N=78	41-50 N=26	> 50 N=17
P95	57	56	57	51	54	59	57	59	56	55	56	-
P90	55	54	55	50	52	57	57	56	55	53	56	54
P85	53	53	53	48	51	56	56	54	53	51	54	52
P80	52	51	52	47	50	55	56	52	52	49	52	51
P75	51	51	50	45	49	54	55	51	51	48	50	50
P70	50	49	50	45	49	53	54	50	50	47	50	48
P65	49	48	49	44	48	52	53	49	49	46	49	47
P60	48	47	48	43	47	51	53	48	48	46	49	46
P55	47	47	48	43	47	50	52	47	48	45	48	44
P50	46	46	47	42	46	50	51	47	47	44	48	44
P45	45	45	46	42	45	49	49	45	46	43	47	43
P40	44	43	45	40	45	48	48	44	45	43	47	42
P35	43	43	44	39	44	47	47	43	44	42	47	41
P30	43	42	43	38	43	46	47	42	43	39	45	39
P25	42	41	42	37	42	46	45	41	42	39	45	39
P20	39	39	40	32	40	45	43	40	41	37	41	38
P15	37	38	37	30	39	43	39	38	38	35	34	38
P10	34	35	33	27	37	41	38	35	34	32	30	35
P5	29	29	28	20	32	32	31	31	29	24	26	28

Términos clasificatorios de la inteligencia propuestos por Raven (1970) adaptados a nuestros resultados	
Grado I	Intelectualmente superior (percentil 95 o más)
Grado II	Claramente por encima de la media (percentil 75-94)
Grado III +	Intelectualmente medio (+) si el percentil está por encima de 50 (51-74)
Grado III	Intelectualmente medio si el percentil está entre los centiles 25 y 75 (50)
Grado III -	Intelectualmente medio (-) si el percentil está por debajo de la mediana (25-49)
Grado IV	Definitivamente inferior a la media si el percentil es inferior a 25 (11-24)
Grado IV (-)	Si el percentil está en o es inferior a 10 (6-10)
Grado V	Deficiente intelectual si el percentil está en 5 o es inferior

IV. Puntuaciones centiles

También hemos querido mostrar a los lectores la tabla de percentiles calculadas en los diferentes grupos constituidos y en la muestra total pues si bien se cuenta con algunas normas estas en general son viejas y no muy representativas, aspecto este que le ha sido criticado al manual del Raven.

Nosotros consideramos que cuando se estudian grupos grandes de sujetos, y estos tienen una cierta

necesidad le presentamos los términos utilizados en la clasificación de inteligencia propuesta por Raven (1970) con el fin de que pueda ser empleada en los datos antes expuestos.

Términos clasificatorios de la inteligencia propuestos por Raven (1970) adaptados a nuestros resultados

Vale aclarar que el experimentador debe, a partir de sus intereses, decidir si utiliza la norma ofrecida

acorde con la escolaridad, edad, sexo o sencillamente emplea la norma de la muestra total.

En todos los casos recomendamos, antes de tomar la decisión, se tenga en cuenta las características (media, desviación estándar, error estándar de la media y número de casos considerados dentro de cada uno de los grupos conformados con el fin de optar por la variante más conveniente acorde a los intereses y/o necesidades particulares que debe satisfacer.

CONCLUSIONES

Establecidos el índice de dificultad y discriminación de la prueba cumplimos con uno de los propósitos que nos habíamos impuesto en el trabajo más ello sólo nos sirve a nuestro juicio como elemento de partida para otras tareas más importantes y, entre ellas, la de proponer una alternativa científicamente probada, que cumpla los requisitos indispensables en cuanto a confiabilidad y validez y que, a su vez, logre ejecutarse en un tiempo más corto lo cual signifique una economía del gasto salarial necesario del experimentador y sus colaboradores tanto en el momento de la aplicación como en el de la calificación. Asimismo, reduzca el tiempo exigido a los sujetos que se someten al proceso de evaluación y/o selección lo cual consideramos también de suma importancia no sólo por su repercusión fisiológica (cuanto mayor duración del tiempo necesario para realizar la prueba mayor fatiga y menor rendimiento), psicológica y, naturalmente, económica.

(Considérese, por ejemplo, que en ocasiones es necesario evaluar miles de sujetos y cada uno de estos pueden devengar un salario medio mensual de alrededor de los \$200.00 lo cual hace que cada hora de trabajo de ellos utilizadas constituya un monto total para todo el grupo bastante elevado).

No todas las variantes propuestas cumplen todos los requisitos antes referidos, sin embargo, el tener la posibilidad de contar con ellas y con estudios de validación de las mismas constituye, a nuestro entender, un paso que contribuya a la solución de un problema actual.

Los ejemplos ilustrativos mostrados creemos que permiten evaluar las posibilidades alternativas del empleo de las dos formas básicas de calificación y, sobre todo, ofrecen al lector la idea necesaria de utilizar uno u otro procedimiento acorde con las necesidades, intereses y/o motivos del proceso de evaluación.

Finalmente, consideramos que los percentiles ofrecidos como resultado de nuestro trabajo cumplen, al menos, todos los requisitos exigidos para su determinación y las estadísticas obtenidas del trabajo efectuado son más amplias que las ofrecidas por el manual del Raven (1970) sobre los contingentes o grupos por ellos referidos.

Lo anterior significa, solamente, que estos son los primeros pasos para la obtención de las verdaderas normas cubanas, reclamo efectuado por numerosos investigadores.

REFERENCIAS

- CAIRO, VALCARCEL, E. (1997): **Estudio de las posibilidades del test de Matrices Progresivas de Raven en una población adulta.** (trabajo inédito).
- _____ (1998): **Estudio del Test de Matrices Progresivas de Raven.** (trabajo inédito).
- CAMPOS PUENTES, A. (s/f). **Experiencias en la aplicación del Test de Matrices Progresivas de Raven para evaluar dirigentes y personal altamente calificado.**
- Comité Ejecutivo del Consejo de Ministros. (s/f). **Bases generales del perfeccionamiento empresarial.** La Habana.
- GARRET, H. (1974): **Estadística en Psicología y Educación.** Ed. Paidós. Buenos Aires.
- KAPLAN, R.M. and D.P. SACUZZO (1985): **Psychological Testing. Principles, applications and Issues.** Brooks/Cole Publishing Company, Pacific Grove, California.
- LIENERT, S.A. (1990): **Selección de lecturas de psicodiagnóstico laboral.** Construcción y análisis de los test. Tomo IV. Edit. ENPES. La Habana.
- PONCE SOLOZABAL, T.L. (1999): **Aproximación al estudio de la validez predictiva del test de Matrices Progresivas de Raven para la selección del personal en el Minint. Trabajo para optar por el título de Master en Psicología Clínica.** La Habana.
- RAVEN, J.C. (1970): **P.M.56, Matrices Progresivas Escala General.** Ed. H.K., Lewis and Co., Ltd./London.
- SANTIESTEBAN REQUENA, C. (1990): **Psicometría. Teoría y práctica en la construcción de tests.** Ediciones Norma S.A. Madrid.