

# Aplicación, corrección y validez de una prueba de rendimiento

Con ser muy laborioso el proceso de planificación y estructuración de una prueba de rendimiento, ya tenga ésta el carácter de global o específica, de profundidad o de rapidez, y aunque se hayan salvado los pasos experimentales previos a su aplicación, de manera que podamos estar seguros de que es adecuada y comprensiva, objetiva y práctica (1), todavía hemos de ser cuidadosos en su aplicación, en su corrección y en la interpretación de los resultados.

Con relación a estas tres últimas operaciones, nunca olvidaremos que siempre que evaluamos lo hacemos persiguiendo un fin o unos propósitos: unas veces intentaremos *apreciar un cambio cualitativo*, logrado mediante el aprendizaje, en la manera o modos de reaccionar ante una situación problemática; otras, quisiéramos convencernos del *rendimiento neto* conseguido en tal período o etapa, ya sea por parte del alumno o en favor de una técnica o del docente. Según sean los objetivos adoptaremos unos procedimientos más o menos rigurosos desde el punto de vista estadístico o para la confiabilidad de la experimentación, pero sin que se menoscabe la seriedad de las medidas y carácter verdaderamente científico.

---

Por AMBROSIO J. PULPILLO RUIZ

---

## EXIGENCIAS PARA UNA BUENA APLICACIÓN DE LAS PRUEBAS DE RENDIMIENTO

Parece obvio advertir que al aplicar cualquier tipo de prueba o examen hayan de tenerse especialmente

(1) Cfr. BUYSE, R.: *La experimentación en Pedagogía*. Edit. Laboral, 1957. págs. 188 a 202.

presentes dos condiciones fundamentales, a saber: la objetividad y la imparcialidad.

La objetividad consiste en actuar evitando todo tipo de prejuicios, imposiciones arbitrarias o dogmáticas y cualquier otro factor de opinión personal. Así, y únicamente así, podremos tener la seguridad de que nuestras mediciones son correctas y que ello nos conducirá a justipreciar bien lo que nos proponíamos examinar. Como se ve, esta condición puede depender tanto del instrumento de evaluación en sí como de su aplicación e interpretación de los resultados a que finalmente lleguemos.

La imparcialidad excluye todo favoritismo y la eliminación del llamado "halo afectivo", de tal manera que, *a priori*, no nos inclinemos a favor de unos examinados o en contra de otros por muchos antecedentes que tengamos de ellos. La opinión hay que dejarla a un lado para que se manifieste la verdad objetiva (2).

Por otra parte, en cualquier tipo de investigación o experimentación, y la evaluación debe tener este carácter siempre, nos encontraremos con factores concomitantes, sistemáticos unos y aleatorios o informales otros, pero siempre actuantes o determinantes.

El buen evaluador debe conocerlos bien para tratar de reducir su influencia al mínimo si es que no es posible neutralizarla o eliminarla por completo. Entre los factores sistemáticos o constantes más decisivos están:

- La validez del instrumento.
- La habilidad o perspicacia del evaluador.
- Las instrucciones que se den al empezar la prueba.
- Las posibilidades que encuentre el examinando de informarse o copiar de otros sobre la marcha de la evaluación.

Los factores fortuitos o aleatorios pueden influir positiva o negativamente y se neutralizan a medida que aumenta la extensión de la muestra o la cantidad de alumnos evaluados. La micro-muestra a este respecto adolece de ese defecto, pero es el caso más corriente en la evaluación de pruebas de rendimiento escolar.

#### LA CORRECCIÓN O Puntuación DE UNA PRUEBA DE RENDIMIENTO

Según sea más o menos grande el número de los *ítems* de la prueba o la cantidad de los alumnos sometidos al proceso de evaluación, adoptaremos

(2) BUYSSE, R.: *Op. cit.*, pág. 125.

para la corrección o el "puntaje" un procedimiento más o menos mecanizado, pudiéndose manejar desde el lápiz y el papel hasta la "plantilla" o cartulina perforada.

Corrientemente se vienen empleando tres formas de calificar la solución de una prueba:

- Negativa, dudosa o positiva.
- Mal, regular, bien y muy bien.
- Otorgando un rango de puntuaciones que va desde 0 a 10.

El primer proceder peca de tanjante y sólo es válido para aquellas ocasiones en que la cuestión se sabe o no se sabe. Es el caso de las pruebas denominadas VERDADERO - FALSO (V - F), y solamente se consiente en calificar también la duda o la omisión (O). Estos *ítems* suelen calificarse con los signos (—), (¿) y (+) en honor a la simplificación.

La segunda forma de calificar una prueba, que es la más generalizada o tradicional, supone las calificaciones de Suspenso, Aprobado, Notable y Sobresaliente, si bien hoy día tales apreciaciones llevan implícitas una puntuación que oscila, como el tercer caso, de 0 a 10.

El traducir cualquier sistema de calificaciones a términos numéricos supone una gran ventaja a la hora de querer llevar los resultados a extremos comparativos o probantes, y la escala de 0 a 10 ofrece la suficiente variedad para dar cabida a las principales matizaciones que debe contener una prueba de rendimiento.

Cuando la prueba es compleja y su valoración total se forma con la suma de escalas de puntuaciones diferentes, para homogeneizar los resultados se suele emplear el procedimiento del "percentilado" (3) o la Escala "M", en virtud de la cual cualquier puntuación se traduce, primero en un valor porcentual que va desde 0,0026 a 99,9969, y luego en el valor "M", comprendido entre 90 y 11 (4).

Digamos, por último, a este respecto que cuando una prueba permite escoger la mejor respuesta entre tres o cinco contestaciones sugeridas, la posibilidad de elegir la correcta supera en algo el conocimiento real que el alumno tiene sobre la cuestión. Una fórmula aceptable a tal efecto es la siguiente (5):

(3) Cfr. GARCÍA HOZ, V.: *Normas elementales de Pedagogía Empírica*. Edit. Escuela Española. Madrid, 1962. 4.ª edic. Págs. 136 a 153.

(4) Cfr. PULPILLO, A.: *Inscripción, Evaluación y Agrupamientos Escolares*. Edit. Escuela Española. Madrid, 1965. 2.ª edic. Págs. 68 a 72.

(5) BUNKER, H.: *Principios fundamentales de evaluación para educadores*. Edit. de la Univ. de Puerto Rico, 1963. Pág. 118.

B = Número de contestaciones buenas o aceptables.  
M = Número de contestaciones malas o erróneas.  
n = Número de alternativas ofrecidas.

$$B = \frac{M}{n-1}$$

que son muy altas a la derecha o a la izquierda (8) y en el fenómeno de la "cabalgadura". El % vencido en el total de la prueba indica si es o no propia de la edad a que se destina.

INTERPRETACIÓN DE DATOS Y VALIDEZ DE UNA PRUEBA DE RENDIMIENTO

Tras la formulación de la prueba, habiendo quedado bien claros los propósitos y los objetivos, y una vez aplicada, sirviéndonos de los instrumentos o técnicas más adecuadas al caso, con su corrección y valoración o puntaje, no queda terminado el proceso de evaluación, sino que todavía falta quizá lo más importante: el análisis y la interpretación de los resultados logrados, que incluyen su ordenación, clasificación y organización, el tratamiento estadístico y la aplicación de criterios para que, una vez convencidos de su validez, podamos formular las conclusiones y recomendaciones que se desprendan real y objetivamente de las mediciones efectuadas.

Con objeto de exponer con la mayor claridad todo esto, vamos a recurrir a un caso práctico. Hace poco más de diez años realizamos en Madrid una experiencia dirigida a medir la validez y fiabilidad de una prueba de instrucción; parte de ella nos viene pintiparada como ejemplo (6).

Se trata del Jalón 2.º de la "Escala sintética de Instrucción" de García Hoz (7), compuesto de 12 ítems, propio para un 2.º curso de escolaridad, niños de siete-ocho años, y que fue sometido a una muestra de 100 alumnos pertenecientes a seis Grupos Escolares distintos. No olvidemos que el valor "probante" de los resultados aumenta con la extensión de la población examinada.

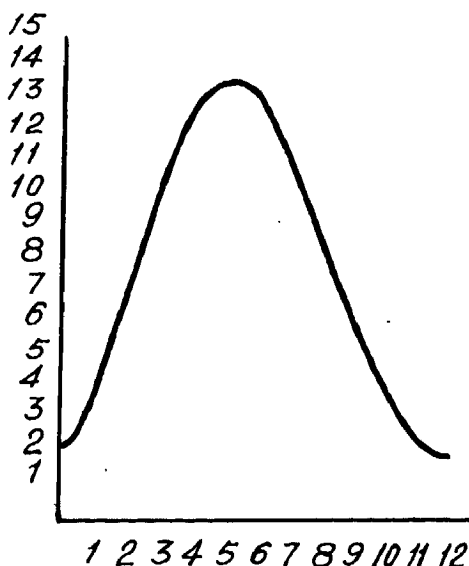
La normalidad de la serie o grupo escogido quedó demostrada al aplicársele con finalidad estadística los criterios de la media, mediana, modo, sigma, error probable, skewness y kurtosis, para asegurarnos de su valor representativo. Como también se sabe que si el grupo de sujetos examinados se ajusta más o menos a la "campana" de Gauss es porque estamos ante una muestra significativa, hemos añadido igualmente la curva de distribución. A estos efectos debe repararse en aquéllas

Puntos	Frecuencias	Criterios
0	0	Me = 5,6
1	6	Md = 5,9
2	6	Mo = 5,0
3	11	S. D = 2,5
4	12	E. P = 1,6
5	16	Sk = 0,360
6	13	Q <sub>1</sub> = 4,16
7	11	Q <sub>3</sub> = 7,0
8	10	P <sub>10</sub> = 2,3
9	7	P <sub>90</sub> = 9,7
10	5	Kur. = 0,191
11	3	
12	0	

Σ de casos: 100

Porcentaje vencido del total de la prueba: 46.6 %.

Curva de frecuencia



(6) Memoria de Licenciatura en Pedagogía del autor de este trabajo, cuyo original obra en la Facultad de Filosofía y Letras de Madrid.  
(7) Cfr. Manual de "tests" para la Escuela. Edit. Escuela Española. Madrid, 1955, págs. 117-126.

(8) Cfr. capítulo de PULPILLO, A., en la obra Orientaciones pedagógicas para directores escolares. Publ. del CEDODEP, Madrid, 1965, pág. 428.

TABLA DE ACIERTOS Y ERRORES DE LA QUE SE DESPRENDE LA DIFICULTAD DE CADA ITEMS (Columna P), ASI COMO EL NIVEL DE CADA SUJETO (Fila T)

Elementos:	1	2	3	4	5	6	7	8	9	10	11	12	T	T <sup>2</sup>
1	-	-	-	-	-	-	-	-	-	-	-	-	0	0
2	-	-	-	-	-	-	-	-	-	-	-	-	0	0
3	-	-	-	-	-	-	-	-	-	-	-	-	0	0
4	-	-	-	-	-	-	-	-	-	-	-	-	0	0
5	-	-	-	-	-	-	-	-	-	-	-	-	0	0
6	-	-	-	-	-	-	-	-	-	-	-	-	0	0
7	-	-	-	-	-	-	-	-	-	-	-	-	0	0
8	-	-	+	-	-	-	-	-	-	-	-	-	1	1
9	-	-	-	-	-	-	-	-	-	+	-	-	1	1
10	-	-	-	-	-	-	-	-	-	-	+	-	1	1
11	+	-	-	-	-	-	-	-	-	-	-	-	1	1
12	-	+	-	-	-	-	-	-	-	-	-	-	1	1
13	+	-	-	-	-	-	-	-	-	-	-	-	1	1
14	-	-	-	-	-	-	+	-	-	-	-	-	1	1
15	-	-	-	+	-	-	-	-	-	-	-	-	1	1
16	-	+	-	-	-	-	-	-	-	-	-	-	1	1
17	-	+	-	-	-	-	+	-	-	-	-	-	2	4
18	-	+	-	-	-	+	-	-	-	-	-	-	2	4
19	-	+	-	-	-	-	-	-	-	-	-	+	2	4
20	-	-	+	-	-	-	+	-	-	-	-	-	2	4
21	-	+	-	-	-	-	-	+	-	-	-	-	2	4
22	-	+	-	-	-	-	-	-	-	-	-	+	2	4
23	-	-	-	-	+	-	-	-	-	-	-	+	2	4
24	+	-	-	-	+	-	-	-	-	-	-	-	2	4
25	-	-	-	-	+	-	-	-	+	-	-	-	2	4
26	+	-	-	-	-	-	+	-	-	-	-	-	2	4
27	+	-	-	-	-	-	-	-	-	+	-	-	2	4
28	-	+	-	-	+	-	-	-	-	-	-	+	3	9
29	-	+	-	-	+	-	-	-	-	-	-	+	3	9
30	+	-	-	-	+	-	+	-	-	-	-	-	3	9
31	+	-	-	-	-	-	+	-	+	-	-	-	3	9
32	-	+	-	-	+	-	+	-	-	-	-	-	3	9
33	-	-	-	-	+	+	-	+	-	-	-	-	3	9
34	+	-	-	-	+	-	+	-	-	-	-	-	3	9
35	-	-	-	-	-	-	-	+	-	+	+	+	3	9
36	-	+	+	+	-	-	-	-	-	-	-	-	3	9
37	+	-	+	-	-	-	-	-	-	-	-	+	3	9
38	-	-	+	-	-	+	-	+	-	-	-	-	3	9
39	-	+	+	-	-	+	-	-	-	-	-	-	3	9
40	-	+	-	-	-	-	-	+	-	+	-	-	3	9
41	+	+	+	-	-	-	-	-	-	-	-	-	4	16
42	+	+	-	-	+	-	-	-	-	-	-	+	4	16
43	+	-	-	-	+	-	+	-	-	-	-	+	4	16
44	+	-	-	-	+	-	+	-	+	-	-	-	4	16
45	+	-	-	-	+	-	+	-	-	-	-	+	4	16
46	+	+	+	-	-	-	+	-	-	-	-	-	4	16
47	-	+	-	-	-	+	-	+	-	-	-	+	4	16
48	+	+	-	-	-	-	+	+	-	-	-	-	4	16
49	+	+	-	-	-	-	+	-	-	-	-	+	4	16
50	+	+	-	-	-	-	+	-	-	-	-	+	4	16
P. parcial:	17	16	11	3	3	16	3	18	1	7	1	14	110	330

Elementos:	1	2	3	4	5	6	7	8	9	10	11	12	T	T <sup>2</sup>
51	+	-	+	+	-	-	-	+	-	-	-	-	4	16
52	+	+	-	-	-	-	-	+	-	-	-	+	4	16
53	+	+	+	-	+	-	-	-	-	-	-	+	5	25
54	-	+	+	-	+	+	-	-	-	-	-	+	5	25
55	+	+	-	-	-	-	+	+	-	-	-	+	5	25
56	+	+	-	-	+	-	-	+	-	-	-	+	5	25
57	+	+	-	-	+	-	-	+	-	-	-	+	5	25
58	+	-	+	+	+	-	-	+	-	-	-	-	5	25
59	-	-	+	+	-	+	-	+	-	-	-	+	5	25
60	+	-	+	-	-	+	-	+	-	-	-	+	5	25
61	-	+	+	-	+	-	-	+	-	-	-	+	5	25
62	+	+	-	-	-	+	-	+	-	-	-	+	5	25
63	+	+	+	-	-	+	-	-	-	-	-	+	5	25
64	-	-	+	+	+	-	-	+	-	-	-	+	5	25
65	+	+	+	-	-	+	-	-	+	+	-	+	6	36
66	+	+	+	+	-	-	-	+	-	+	-	+	6	36
67	-	+	+	-	+	+	-	+	-	-	-	+	6	36
68	-	+	+	-	+	+	+	-	-	-	-	+	6	36
69	+	-	+	+	-	-	-	+	+	-	-	+	6	36
70	+	-	-	-	-	+	+	+	-	-	-	+	6	36
71	+	+	-	+	-	+	+	+	-	-	-	-	6	36
72	+	-	+	+	-	-	-	+	-	-	-	+	6	36
73	-	+	+	+	-	+	+	+	-	-	-	-	6	36
74	+	+	-	-	+	-	+	-	+	-	-	+	6	36
75	-	+	+	+	+	-	+	+	-	-	-	+	7	49
76	+	+	+	+	-	+	-	+	-	-	-	+	7	49
77	+	+	+	+	-	-	-	+	-	-	-	+	7	49
78	+	+	-	-	+	+	-	+	+	-	-	+	7	49
79	+	+	-	+	+	-	-	+	+	-	-	+	7	49
80	+	+	+	-	+	+	+	+	-	-	-	-	7	49
81	+	+	-	-	+	-	+	+	-	-	-	+	7	49
82	+	+	-	-	+	-	+	+	+	-	-	+	7	49
83	+	+	-	-	+	-	+	+	+	-	-	+	7	49
84	+	+	+	+	+	+	+	-	-	-	-	+	8	64
85	+	+	+	+	+	+	-	+	-	-	-	+	8	64
86	+	+	+	+	-	+	-	+	+	-	-	+	8	64
87	+	-	+	+	-	+	+	+	-	-	-	+	8	64
88	+	+	-	-	+	+	+	+	-	+	-	-	8	64
89	+	+	+	+	-	+	-	+	+	-	-	+	8	64
90	+	+	+	+	+	+	+	+	-	-	-	+	8	64
91	+	+	+	+	+	+	-	+	-	-	-	+	8	64
92	+	+	-	+	+	+	-	+	+	-	-	+	8	64
93	+	+	+	+	+	+	+	+	-	-	-	+	9	81
94	+	+	+	+	+	-	+	+	+	-	-	+	9	81
95	+	+	+	+	+	+	-	+	+	-	-	+	9	81
96	+	+	+	+	-	+	+	+	-	+	-	+	9	81
97	+	+	+	+	+	+	-	+	-	+	-	+	9	81
98	+	+	+	+	+	+	+	+	-	+	-	+	10	100
99	+	+	+	+	+	+	+	+	+	-	+	+	10	100
100	+	+	-	+	+	+	+	+	+	+	+	+	11	121
P. total:	59	57	43	32	30	49	20	61	17	15	7	58	449	2.765

$$P^2 = \left\{ \begin{matrix} 3.481 \\ 3.249 \\ 1.849 \\ 1.024 \\ 900 \\ 2.401 \\ 400 \\ 3.721 \\ 289 \\ 225 \\ 49 \\ 3.364 \end{matrix} \right\} = 20.951$$

$$449^2 = 201.601$$

Indice comparativo de dificultad y discriminación global (% de aciertos × % de errores) y analítica o por grados. Los alumnos que vencieron 1/3 de los ítems forman el grado 1.º o A; los que alcanzaron los 2/3 son los del 2.º grado o B, y los que rebasaron los 2/3 se consideraron del grado superior o C. (Si una prueba separa claramente al alumno brillante del normal y a éste del torpe, aumenta considerablemente su fiabilidad.)

Elementos	Dificultad	Dis. glob.	Disc. por grados		
			G. A.	G. B.	G. C.
1	25	1.875	43	90	100
2	64	2.304	23	36	67
3	97	291	0	4	7
4	13	1.131	69	96	100
-5	32	2.176	37	80	100
6	62	2.356	11	44	80
7	67	2.211	14	28	93
8	49	2.499	9	68	93
9	65	2.275	3	44	80
10	79	1.659	14	12	67
11	30	2.100	37	84	100
12	57	2.451	23	44	87

El índice de selectividad de cada ítem, siguiendo el método de Hort, pone de manifiesto aquellos de valor negativo o nulo que afectan a la consistencia interna de la prueba en el sentido de que en nada le favorecen o perjudican. Veamos a continuación la tabla de coincidencias:

	1	2	3	4	5	6	7	8	9	10	11	12	Sumas
1	75	29	2	70	61	32	27	43	33	14	57	32	476
2		36	2	36	27	14	13	20	14	9	27	16	243
3			3	3	3	2	2	1	1	0	2	2	23
4				87	68	30	31	50	33	18	63	36	525
5					68	29	25	40	30	16	50	28	445
6						38	22	27	19	10	30	22	275
7							33	25	15	12	26	18	249
8								51	22	12	43	27	361
9									35	11	31	23	267
10										21	17	15	155
11											70	37	453
12												43	299

De la anterior tabla pasamos a los siguientes cálculos:

Para las coovarianzas de cada ítem:

$$\frac{1}{N} \left[ \Sigma \text{ de Coincidencias} - \left( \frac{\text{Puntos totales del ítem}}{N} \times \text{Puntos totales de la prueba} \right) \right]$$

En el ítem 1:

$$\frac{1}{100} \left[ 476 - \left( \frac{75}{100} \times 560 \right) \right] = 0,5600$$

Para las varianzas de cada ítem:

$$\frac{\text{Puntos totales}}{N} - \left( \frac{\text{Puntos totales}}{N} \right)^2$$

En el ítem 1:

$$\frac{75}{100} - \left( \frac{75}{100} \right)^2 = 0,1875$$

La  $\Sigma$  de Coovarianzas es igual a la diferencia entre la coovarianza y la varianza de cada ítem. En el 1:

$$0,5600 - 0,1875 = 0,3725$$

El índice de selección es igual a  $\Sigma$  de Coovarianzas — Varianza del ítem. En el 1:

$$0,3725 - 0,1875 = 0,1850$$

### MATRIZ DE COOVARIANZAS Y SELECTIVIDAD

Elementos	$\Sigma$ de coincid.	$\Sigma$ Coovarianzas de elementos	Varianzas de cad. elem.	De las coovarianzas	Índice de selección
1	476	0,5600	0,1875	0,3725	0,1850
2	243	0,4140	0,2304	0,1836	0,0468
3	23	0,0620	0,0891	0,0271	0,1162
4	525	0,3780	0,1131	0,2649	0,1518
5	445	0,6420	0,2176	0,4244	0,2068
6	275	0,6220	0,2356	0,3864	0,1508
7	249	0,6420	0,2211	0,4209	0,1998
8	361	0,7540	0,2499	0,5041	0,2542
9	267	0,6100	0,2275	0,3825	0,0550
10	155	0,3740	0,1659	0,2081	0,0422
11	453	0,6100	0,2100	0,4000	0,1900
12	299	0,5820	0,2451	0,3369	0,0918

Para la obtención de coeficiente de fiabilidad de la prueba recurrimos al método de Hort (9), previo el análisis de varianza consiguiente:

Origen de varianza	$\Sigma$ cuadrat.	G. de libertad	Cuadrados 1/2
Entre sujetos ...	54,83	99	0,55
Entre elementos.	65,39	11	5,94
Residual o de error... ..	178,44	1.089	0,16
<b>TOTALES ...</b>	<b>298,66</b>	<b>1.199</b>	

$$r_{tt} = \frac{0,55 - 0,16}{0,55} = 0,70$$

#### FORMULAS DE OBTENCION

$$\Sigma \text{ de cuadrados entre sujetos: } \frac{\Sigma T^2}{n} - \frac{(\Sigma T)^2}{nN} = \frac{3.794}{12} - \frac{313.600}{1.200} = 54,83$$

$$\Sigma \text{ de cuadrados entre ítems: } \frac{\Sigma P^2}{n} - \frac{(\Sigma T)^2}{nN} = \frac{32.672}{100} - \frac{313.600}{1.200} = 65,39$$

$$\Sigma \text{ total de cuadrados: } \frac{\Sigma T(nN - \Sigma T)}{nN} = \frac{560(1.200 - 560)}{1.200} = 298,66$$

$\Sigma$  residual o de errores:

$$\Sigma \text{ total de cuadrados} - (\Sigma \text{ de cuadrados entre sujetos} + \Sigma \text{ de cuadrados entre ítems}) = 298,66 - (54,83 + 65,39) = 178,44$$

$$\text{Grados de libertad: } N - 1 = 99; \quad n - 1 = 11; \\ (N - 1)(n - 1)^2 = 1.089$$

Los cuadros medios vienen dados por el cociente entre  $\Sigma$  de cuadrados y grados de libertad correspondiente.

(9) Todas las fórmulas y métodos descritos fueron tomados y estudiados en las clases de Didáctica de la Sección de Pedagogía de la Facultad de Filosofía y Letras profesadas por el Dr. Fernández-Huerta.

Con todo ello llegamos, por fin, al índice de fiabilidad de la prueba:

$$r = \frac{\text{Cuadrado medio entre sujetos} - \text{Cuadrado residual}}{\text{Cuadrado medio entre sujetos}} = \frac{0,55 - 0,16}{0,55} = 0,70$$

#### CONCLUSIONES

1.ª En general, podemos decir que la prueba que ha sido objeto de evaluación ha recaído sobre un grupo representativo y normal, como lo indica la curva de frecuencias. La media real de las puntuaciones obtenidas (5,6) está muy próxima a la media teórica (6) y también a la mediana (5,9) y al modo (5,0). Los restantes criterios tampoco se apartan mucho de la normalidad.

2.ª La prueba en sí ha resultado difícil, puesto que no llegaron a superarla nada más que el 46,6 por 100 de los examinados, y los ítems considerados aisladamente presentan algunos, como el 3 y el 10, un alto índice de dificultad. En cambio, el 4 es el más fácil de todos.

3.ª Salvo el ítem 3, de escasa discriminación global y gradual, y el 10, que no diferencia bien a los alumnos flojos de los medianos, los restantes elementos de la prueba son aceptables desde este punto de vista.

4.ª El ítem 3, además de su escasa discriminación, presenta, como el 2, un índice de selección negativo, lo que perjudica la consistencia interna de la prueba, motivo suficiente para ser retirados ambos de ella.

5.ª La varianza entre los ítems es un poco alta; sin embargo, el índice de fiabilidad total de la prueba es bueno (0,70).

6.ª Un reajuste de la prueba en su conjunto, para hacerla más fácil, y la sustitución por otros de los elementos discordantes, parece aconsejable si se quiere que a su valoración como prueba de control de rendimiento se una el de indicativa para la promoción.

Hubiera deseado el autor de este trabajo que determinadas alusiones estadísticas, otras referidas a métodos y fórmulas, quedasen más ampliamente expuestas, así como también haber reproducido con detalle la prueba objeto de valoración con la descripción de sus singularidades intrínsecas y aquellas otras manifestadas en su constatación, pero todo ello quizá hubiera hecho más árido el estudio y, sobre todo, necesitaba un espacio mucho mayor del que hemos dispuesto.

Quede, pues, como orientación y referencia para todos aquellos educadores que están interesados por estas cuestiones, y como prueba de que la construcción, aplicación y validez de una prueba de rendimiento exige hondura en su tratamiento y seriedad en el análisis de sus resultados.