

## Ejercicios

**5.1.** ¿Qué tipo de evidencia de validez se persigue al aplicar los siguientes procedimientos?

- Establecer la relación entre cada ítem de un examen y el objetivo instruccional de un curso.
- Correlacionar las puntuaciones obtenidas en un test de habilidades mecánicas con la evaluación de la ejecución de una tarea de carpintería.
- Correlacionar la forma corta y la forma larga de un mismo test.

**5.2.** Un psicólogo construye una escala para medir la *Necesidad de Cierre Cognitivo* (NCC). Este constructo se refiere a la motivación de las personas por buscar y mantener una respuesta definitiva ante un problema determinado. Además de estudiar la fiabilidad de las puntuaciones obtenidas, plantea varios trabajos para obtener evidencias sobre la validez de las mismas; En uno de esos trabajos propone estudiar la relación entre las puntuaciones en la escala NCC y otras dos escalas. Una de ellas mide la Necesidad de cognición (NC) que refiere a la preferencia que muestran las personas hacia la actividad de pensar y la otra la Necesidad de Evaluación (NE), la tendencia de las personas a juzgar los objetos, las ideas y a las personas en términos evaluativos (ej., bueno-malo). El autor eligió analizar las relaciones de la NCC con la NC y la NE porque estos tres constructos comparten una misma naturaleza metacognitiva en sus medidas. Es decir, en los tres casos las personas tienen que pensar sobre su propia actividad de pensar. En cuanto a la relación esperada entre NCC y NC, pronosticaba que las puntuaciones de los participantes NCC correlacionarían de forma negativa con sus puntuaciones en NC, puesto que las personas a las que les gusta pensar no suelen buscar una respuesta rápida y definitiva para dejar de pensar sobre un problema, sino todo lo contrario. En cuanto a la relación entre la NCC y la NE, consideraba que la relación podría ser más abierta. Por un lado, se podría pronosticar una correlación moderadamente positiva puesto que la evaluación en términos de bueno o malo ofrece una respuesta inmediata. Por otro lado, la NCC y la NE podrían no correlacionar o incluso mostrar una correlación moderadamente negativa, ya que las respuestas evaluativas a veces no se perciben tan definitivas como otro tipo de juicios.

- ¿Qué fuente de evidencia sobre la validez buscaba el autor al estudiar la relación entre estos tres constructos (NCC, NC y NE)?
- Determine si la decisión de buscar evidencias sobre la validez de NCC examinando su correlación con cada una de las dos escalas (NC y NE) es o no acertada. Argumente su respuesta.

**5.3.** Se pidió a un grupo de 4 jueces expertos que evaluaran el grado en que 4 ítems de un examen de Matemáticas medían cada uno de los 5 componentes que el examen deseaba evaluar: Álgebra (A), Geometría (G), Cálculo (C), Probabilidad (P) y Medición (M) con el objetivo de obtener el índice de congruencia de Rovinelli y Hambleton.

- La Tabla 1 muestra dos ítems<sup>1</sup>. El ítem 1 está diseñado para medir solamente Álgebra y el ítem 2 para medir solamente Geometría. Complimente la tabla con el patrón ideal de evaluaciones que un juez daría a estos dos ítems si se cumplieran las especificaciones para las que fueron diseñados:

Tabla 1. Dos ítems del examen de Matemáticas.

Ítems	A	G	C	P	M
Ítem 1. Calcule $x$ : $3x + 2 = 5$					
Ítem 2. ¿Cuál es la amplitud de un ángulo recto?					

- ¿Consideraría usted adecuado utilizar el Índice de Congruencia de Rovinelli y Hambleton para evaluar la validez de contenido de un ítem que haya sido diseñado para evaluar más de un objetivo instruccional? Razone su respuesta.
- A partir de los datos de la Tabla 2, obtenga las evaluaciones medias obtenidas por los 4 ítems en los 5 objetivos instruccionales medidos.
- Obtenga e interprete el índice de Rovinelli y Hambleton para cada uno de los ítems de la Tabla 2.

<sup>1</sup> Este ejemplo ha sido adaptado de Turner, R.C.; Mulvenon, S. W.; Thomas S. P. & Balkin, R. S. (2002, April). Computing Indices of Item Congruence for Test Development Validity Assessments. Proceedings. SUGI 27. Paper 255-27. Orlando, Florida.

Tabla 2. Evaluaciones de 4 jueces en 4 ítems y 5 posibles objetivos

Jueces	Ítems	Objetivos instruccionales				
		1	2	3	4	5
1	1	+1	-1	-1	-1	-1
1	2	+1	0	-1	-1	-1
1	3	-1	-1	+1	-1	-1
1	4	-1	-1	+1	-1	-1
2	1	+1	-1	-1	-1	-1
2	2	+1	-1	-1	-1	-1
2	3	-1	-1	+1	-1	-1
2	4	-1	-1	+1	-1	-1
3	1	+1	-1	-1	-1	-1
3	2	+1	-1	-1	-1	-1
3	3	-1	-1	+1	-1	-1
3	4	-1	-1	+1	-1	-1
4	1	+1	-1	-1	-1	-1
4	2	+1	-1	-1	-1	-1
4	3	+1	-1	-1	-1	-1
4	4	+1	-1	0	-1	-1

5.4. A continuación se ofrece una matriz multirasgo-multimétodo<sup>2</sup> incompleta que ofrece las correlaciones obtenidas al evaluar las actitudes hacia figuras con autoridad y figuras de iguales utilizando distintos métodos.

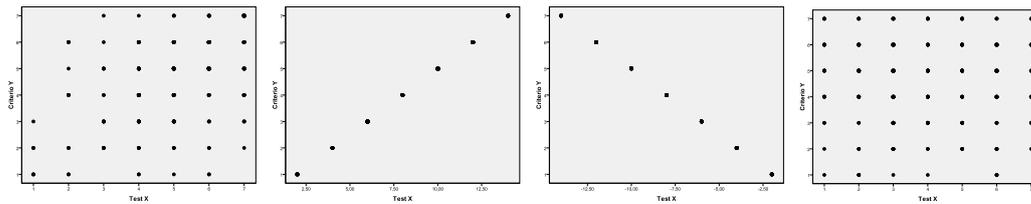
	Entrevista			Lista de adjetivos		
	Padre	Jefe	Compañero	Padre	Jefe	Compañero
Entrevista	Padre					
	Jefe	0,64				
	Compañero	0,65	0,76			
Lista de adjetivos	Padre	0,40	0,08	0,09	0,24	
	Jefe	0,19	0,10	0,03	0,23	0,34
	Compañero	0,27	0,11	0,23	0,21	0,45

- ¿Cuántos métodos incluye?
- ¿Cuántos rasgos mide?
- Identifique, poniendo entre paréntesis sus valores, e interprete las correlaciones monorrasgo-monométodo.
- Identifique, poniendo en negrita sus valores, e interprete los coeficientes que conforman la *diagonal de la validez*.
- Identifique, subrayando sus valores, e interprete las correlaciones heterorrasgo-monométodo.
- ¿Cómo se denominan el resto de valores de la matriz no identificados? Interprete su significado.

5.5. En el siguiente ejercicio se muestran 4 diagramas de dispersión correspondientes a las puntuaciones de 311 personas en un test  $X$  y un criterio  $Y$ . Se proporcionan además 4 valores de coeficientes de validez: 1, 0,4, 0, y -1. Indique qué coeficiente de validez corresponde a cada uno de los diagramas de dispersión.

<sup>2</sup>adaptada de Campbell y Fiske (1959)

Grafico 1 Grafico 2 Gráfico 3 Gráfico 4



**5.6.** Si el coeficiente de validez de un test respecto a un determinado criterio ( $Y_1$ ) es el doble del obtenido respecto a un segundo criterio ( $Y_2$ ). ¿Cuánto mayor será entonces el coeficiente de determinación del test respecto al primer criterio comparado con el obtenido respecto al segundo criterio?

- a) Será dos veces mayor
- b) Será cuatro veces mayor
- c) Será ocho veces mayor

**5.7.** A continuación se detallan las puntuaciones que 10 personas obtuvieron en un test de rendimiento escolar ( $X$ ) y las calificaciones medias del curso ( $Y$ ), que se consideran como un criterio de aprovechamiento académico.

Personas	1	2	3	4	5	6	7	8	9	10
Test	8	15	2	10	8	4	5	12	9	3
Calificación	9	8	7	6	4	2	4	4	5	2

- a) Suponiendo que las 10 personas constituyen un grupo normativo apropiado, obtenga e interprete el coeficiente de validez del test.
- b) Obtenga el error típico de estimación del test

**5.8.** En la selección de aspirantes a un curso de formación, los sujetos han sido examinados con un cuestionario que obtuvo una media de 5 y una desviación típica de 2. Una vez terminado el curso, los mismos sujetos fueron valorados por sus formadores según una escala de 0 a 20, con media 10 y desviación típica 3. La correlación entre los resultados en el cuestionario y las valoraciones de los formadores fue de 0,35.

- a) Realice una estimación puntual de la valoración que recibiría una persona que obtuvo 4 puntos en el cuestionario.
- b) Estime, con probabilidad 0,95, el intervalo de confianza en el que se encontrará la valoración para esta misma persona.

**5.9.** El coeficiente de determinación de un test es 0,25 y la varianza del criterio es 2. Obtenga:

- a) El coeficiente de validez y la varianza de los errores de pronóstico.
- b) La amplitud que tendrá la estimación por intervalo en el criterio para cualquier persona, con probabilidad 0,95.

**5.10.** Aplicamos a Juana un test para predecir su rendimiento en un criterio, concluyendo que, con probabilidad 0,99, su puntuación en dicho criterio estará entre 24 y 30. Responda razonadamente a las siguientes cuestiones:

- a) ¿Cuál es la estimación puntual que hemos realizado a Juana?
- b) Si hubiéramos establecido el intervalo con probabilidad 0,95 ¿Cómo afectaría esto a la amplitud del intervalo?

**5.11.** Indique si las siguientes afirmaciones referidas a la validez de las puntuaciones de un test son verdaderas o falsas. Razone su respuesta.

- a) La validez de constructo de las puntuaciones asegura su validez referida a un criterio.
- b) Para estudiar la validez de contenido de un test **no** es necesario aplicarlo a una muestra.
- c) Si el coeficiente de validez de un test vale 0,5 esto significa que el 50% de la variabilidad de las puntuaciones en el criterio se puede pronosticar a partir del test.

**5.12.** En un test con 25 ítems el coeficiente de fiabilidad es 0,25 y el de validez es 0,10. Si correlacionamos las puntuaciones del test con otro criterio.

- a) ¿Podríamos obtener un coeficiente de validez menor de 0,10?  
b) ¿Podríamos obtener un coeficiente de validez mayor de 0,60?

**5.13.** Si el coeficiente de validez de un test es 0,37, entonces la correlación entre las puntuaciones verdaderas de dicho test y el criterio será...

- a) mayor que 0,37  
b) Igual a 0,37  
c) menor que 0,37

**5.14.** Sean 0,34 y 0,19 los coeficientes de validez obtenidos al correlacionar las puntuaciones en un test y un criterio. Uno de ellos se obtuvo utilizando todas las puntuaciones de la muestra y el otro utilizando el 40% de la muestra que se correspondía con las puntuaciones más altas en el test ¿Cuál de ellos se obtuvo con la muestra total? Razone su respuesta.

**5.15.** Una empresa dedicada a la selección de personal recibe el encargo de confeccionar un test que permita seleccionar trabajadores que probablemente permanezcan en la empresa durante al menos 3 años. Un psicólogo construye el test y pide a la mencionada empresa que lo aplique a sus trabajadores. La empresa decide aplicar el test solo a aquellos trabajadores que llevaban con ellos más de 4 años, por considerarles los más aptos. El bajo valor obtenido para el coeficiente de validez (se muestra en la tabla siguiente) provocó numerosas críticas al trabajo del psicólogo ¿Se le ocurre algún argumento que usted como psicólogo podría utilizar para defender la validez de la prueba construida?

<i>N</i>	<i>Coficiente alfa del test</i>	<i>Coficiente de validez</i>	<i>Media de "años de permanencia en la empresa"</i>	<i>Varianza del test</i>	<i>Varianza de "años de permanencia en la empresa"</i>
345	0,82	0,23	6,7	42,31	0,04

**5.16.** Sabemos que aumentando la longitud de un test, podemos aumentar también su fiabilidad, y que la fiabilidad del test es un factor que permite incrementar la validez del test. Queremos obtener un coeficiente de validez de 0,8 ( $R_{XY}$ ) y sabemos que la fiabilidad del test es 0,8 ( $r_{XX}$ ) y la del criterio es 0,6 ( $r_{YY}$ ). ¿Lograremos nuestro objetivo aumentando la fiabilidad del test?

**5.17.** Un psicólogo social diseña un test con 5 ítems y obtiene los coeficientes de fiabilidad, de 0,4, y validez, de 0,36. En vista de estos valores tan bajos, decide rechazar el test. Valore si la decisión del psicólogo de rechazar el test a causa de sus bajos coeficientes de fiabilidad y validez es la más adecuada.

**5.18.** Un test de 5 ítems tiene un coeficiente de fiabilidad de 0,5 y un coeficiente de validez de 0,4

- a) Queremos que su coeficiente de validez alcance el valor de 0,5 ¿Qué longitud debería tener el test?  
b) Queremos que su coeficiente de validez alcance el valor de 0,6 ¿Qué longitud debería tener el test?  
c) ¿Cuál es máximo valor del coeficiente de validez que se puede alcanzar alargando el test?  
d) ¿Cuánto valdría el porcentaje de varianza del test explicado por el criterio si añadiésemos 2 formas paralelas.