



# Psicometría

## Tema 6 VALIDEZ DE LAS INFERENCIAS I

Psicometría

**M<sup>a</sup> Isabel García Barbero**, UNED. Universidad  
Nacional de Educación a Distancia, 2006

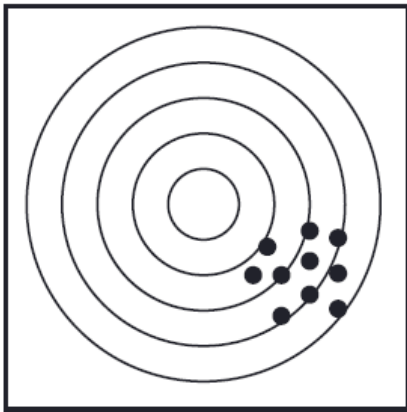
Esperanza Sánchez Moya



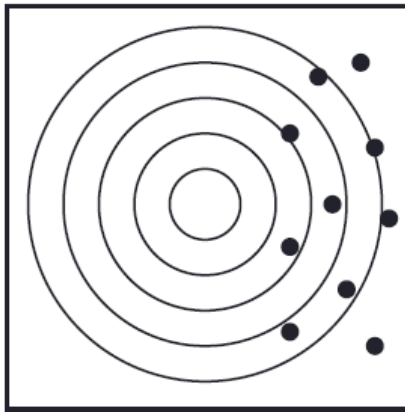
- La **fiabilidad** de un instrumento de medida es una **condición necesaria pero no suficiente**
- Debemos plantearnos hasta qué punto el instrumento de medida sirve para la finalidad perseguida con su aplicación
- Los instrumentos de medida deben proporcionar medidas **fiables** a partir de las cuales se pueden hacer inferencias **validas**
- **VALIDEZ:** Hace referencia al grado en que el test mide aquello que pretende medir
- Con **validez** nos referimos a **una determinada interpretación de las puntuaciones**
  - \* Lo que se valida no es el instrumento sino, como señala Corbacho (1971) la **interpretación** de los datos obtenidos por medio de un procedimiento específico .
- **Fiabilidad y validez** no son propiedades intrínsecas del test sino de los **objetivos de su uso y fines**. Se puede decir que la validez no se refiere al test en si mismo sino al buen uso de los resultados



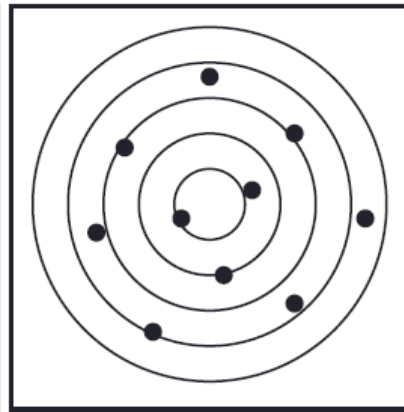
*Validez.*



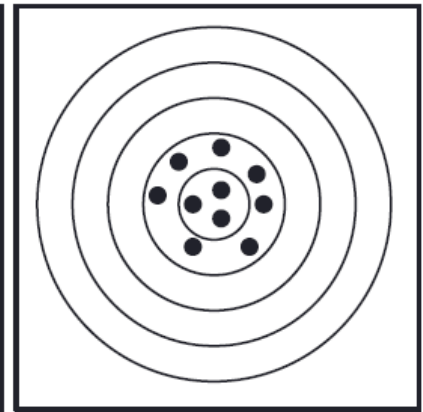
Alta fiabilidad  
Baja validez



Baja fiabilidad  
Baja validez



Baja fiabilidad  
¿Alta validez?



Alta fiabilidad  
Alta validez

Fuente: Tema 6 Psicometría. Universidad Camilo José Cela



# Tipos de VALIDEZ

dependiendo de los **Objetivos** del test (API 1966)

## Contenido

- Para determinar el rendimiento o actuación de un sujeto en un universo de situaciones
- Relevancia y Representatividad de los ítems

## Constructo

- Para inferir el grado en que un sujeto posee algún rasgo atributo
- Medidas del mismo constructo obtenidas con otras prueba

## Correlacional

- **Concurrente**: test y criterio simultáneamente (diagnostico)
- **Predictiva** : criterio con posterioridad al test (selección y clasificación)
- **Retrospectiva**: criterio con anterioridad al test



# Validez de Contenido

- Analiza e indica si los ítems o elementos del test son una **muestra relevante y representativa del constructo** sobre el que se va hacer inferencias.
  
- La forma típica para llevar a cabo un estudio de validez de contenido es utilizar un **grupo de expertos** que analizaran dos aspectos:
  - ✓ Que el test NO incluya aspectos irrelevantes del dominio de interés
  - ✓ Que el tests incluya todos los elementos importantes que definen el dominio
  
- Para llevar a cabo la especificación del domino hay que:
  - ✓ Analizar las áreas de contenido que se quieren cubrir
  - ✓ Se deben analizar los proceso que se van a evaluar y la importancia relativa de cada uno de ellos



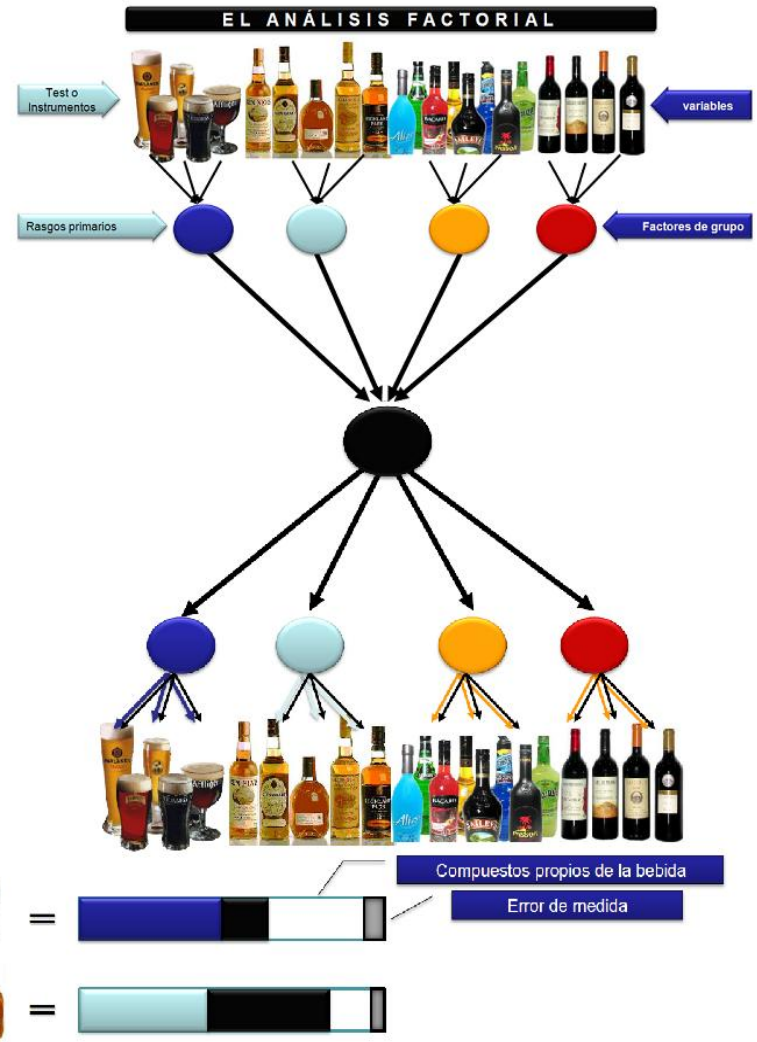
# Validez de Constructo

- Da **significado** a las puntuaciones de los test, pues **permite obtener evidencia** de que las conductas observables que se han elegido como indicadores del constructo (variable latente inobservable) realmente lo son
  
- Se centran en el análisis de las relaciones entre:
  - ✓ Las puntuaciones en los distintos ítems (análisis de la estructura interna del test)
  - ✓ Las relaciones entre el test y otras medidas del mismo constructo (estructura externa del test).
  
- Para ello debemos
  - ✓ **Definir** el constructo y su hipotética relación con determinadas conductas y con otros constructos.
  - ✓ **Diseñar** el instrumento, seleccionando ítems relevantes y representativos del constructo
  - ✓ **Obtener** datos empíricos de las relaciones entre las puntuaciones del test y las conductas observables.



# 1. Análisis Factorial

- Nos permite descubrir la estructura que subyace a las puntuaciones obtenidas por los sujetos en los distintos ítems del test o en un conjunto de test
- Es un conjunto de técnicas estadísticas empleadas para resumir un conjunto de variables observadas en un menor número de variables latentes (no observadas).
- Sirve para validar la estructura interna de una prueba. Analiza si hay una dimensión o varias en el test
- Dos enfoques: confirmatorio y exploratorio
- Desde el enfoque confirmativo se establece a priori hipótesis acerca de la estructura subyacente y del número de dimensiones, pero no pasa así en el enfoque exploratorio



Ejemplo: Los marcianos se van de juerga (Fuente: Sergio Escorial, UCM)



## 2. Matriz multimetodo-multirrasgo MMMR

Medir un mismo constructo mediante distintos procedimientos y distintos constructos mediante el mismo procedimiento

- **VALIDEZ CONVERGENTE (VC)**
  - ✓ Correlaciones entre las medidas obtenidas del **mismo constructo** a través de **distintos procedimientos**
  - ✓ **Mismo rasgo distinto método:** mono rasgo - heterometodo
  
- **VALIDEZ DISCRIMINANTE (VD)**
  - ✓ Correlaciones entre las medidas obtenidas de **distintos constructos** con el **mismo procedimiento**
  - ✓ **Distinto rasgo mismo método:** heterorasgo - monometodo
  - ✓ La **VD < VC**
  
- **FIABILIDAD**
  - ✓ **Mismo rasgo mismo método**, heterorasgo-heterometodo relación entre cada constructo medido con el mismo método (diagonal de la matriz)





# Matriz Multimétodo Multirrasgo - MMMR

Matriz de correlaciones en la que 3 rasgos: Extraversión, Liderazgo e Inteligencia Social se midieron con 3 métodos distintos: AI Autoinforme, OS Observación Sistemática y EP Encuesta a Profesores

		Extraversión			Liderazgo			Inteligencia social		
		AI	OS	EP	AI	OS	EP	AI	OS	EP
Extraversión	AI	0.89								
	OS	0.51	0.89							
	EP	0.38	0.37	0.76						
Liderazgo	AI	0.57	0.22	0.09	0.93					
	OS	0.22	0.57	0.10	0.68	0.94				
	EP	0.11	0.11	0.46	0.59	0.58	0.84			
Inteligencia social	AI	0.56	0.22	0.11	0.67	0.42	0.33	0.94		
	OS	0.23	0.58	0.12	0.43	0.66	0.34	0.67	0.92	
	EP	0.11	0.11	0.45	0.34	0.32	0.58	0.58	0.60	0.85



### Ejemplo:

Una empresa desea medir el grado de sociabilidad y la ansiedad de sus directivos. Para medir cada uno de estos dos rasgos se dispone de dos tipos de tests: un test de verdadero-falso de aplicación colectiva (VF) y otro manipulativo de aplicación individual (MI). Los cuatro tests se aplican a una misma muestra y la matriz multirrasgo Multimétodo resultante es la siguiente:

- a) ¿ Qué se puede decir de la validez convergente de dichos tests?
- b) ¿ Qué se puede decir de su validez discriminativa?

	Sociabilidad V-F	Ansiedad V-F	Sociabilidad MI	Ansiedad MI
Sociabilidad V-F	0.90	0.20	0.85	0.20
Ansiedad V-F	0.20	0.89	0.10	0.82
Sociabilidad MI	0.85	0.10	0.92	0.18
Ansiedad MI	0.20	0.82	0.18	0.87



## Ejemplo del libro

**124.** A continuación se presentan los resultados de un estudio Multirrasgo-Multimétodo donde se ha medido la correlación entre la introversión (I), la agresividad (A) y la sociabilidad (S) mediante dos métodos diferentes (método A y B):

		Método A			Método B		
		I	A	S	I	A	S
Método A	I	0,96					
	A	0,29	0,85				
	S	0,56	0,40	0,93			
Método B	I	0,80	0,33	0,58	0,96		
	A	0,20	0,65	0,44	0,40	0,80	
	S	0,30	0,31	0,63	0,56	0,25	0,20

Indicar los coeficientes de fiabilidad, validez convergente y discriminante.



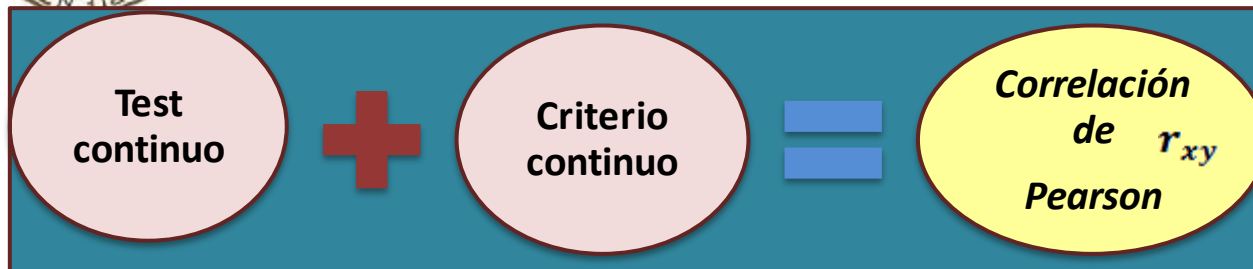
# Validación referida al Criterio

- El objetivo principal es evaluar la hipótesis de **relación entre el test y criterio**
- Se utilizan 2 tipos de índices o medidas para describir la capacidad de un test para predecir un criterio:
  - ✓ Medidas correlacionales
  - ✓ Medidas de error en la predicción
- **Elección de los indicadores:**
  - ✓ Que sean relevantes
  - ✓ Que estén libres de sesgo
  - ✓ Que sean fiables
  - ✓ Deben ser accesibles
- **2 tipos de Diseño**
  - ✓ Un **único test predictor y un solo indicador del criterio**: correlación y el modelo de regresión lineal simple
  - ✓ Varios **predictores y un solo indicador del criterio**: correlación y regresión lineal múltiple



# Coeficiente de Validez: correlación test (X) y criterio (Y)

diferentes coeficientes dependiendo del nivel de medida del test del criterio



$$r_{xy} = \frac{n \sum XY - \sum X \sum Y}{\sqrt{n \sum X^2 - (\sum X)^2} \sqrt{n \sum Y^2 - (\sum Y)^2}}$$

$$r_{bp} = \frac{\bar{X}_p - \bar{X}}{S_x} \sqrt{p/q}$$



$$r_b = \frac{\bar{X}_p - \bar{X}}{S_x} p/y$$



		X	
		0	1
Y	0	a	b
	1	c	d

$$\phi = \frac{bc - ad}{\sqrt{(a + b)(c + d)(a + c)(b + d)}}$$



$$\phi_b = \frac{bc - ad}{\sqrt{(a + b)(c + d)}} \frac{\sqrt{p q}}{y}$$





## Ejemplo

Se ha desarrollado un nuevo test para evaluar la satisfacción laboral. En la tabla adjunta se presentan los resultados obtenidos por 8 sujetos en dicho test, así como sus propias valoraciones (criterio) emitidas 1 año antes. Calcular su coeficiente de validez.

Sujetos	Criterio Y	Test X
A	1	2
B	2	4
C	3	5
D	4	6
E	5	3
F	6	8
G	7	7
H	8	8



# Rectas de Regresión

Sirven para **PRONOSTICAR O PREDECIR** como se comportara la variable criterio tomando en consideración la variable predictora X

	PUNTUACIÓN	RECTA REGRESIÓN	SIGNIFICADO	CALCULO
DIRECTAS	X	$Y' = A + B X$	<b>A</b> ordenada en el origen o intercepto  <b>B</b> pendiente, coeficiente de regresión o tasa de cambio	$Y' = r_{xy} \frac{S_y}{S_x} (X - \bar{X}) + \bar{Y}$
DIFERENCIALES	$x = X - \bar{X}$	$y' = b x$	$b = B$ pendiente	$y' = r_{xy} \frac{S_y}{S_x} (X - \bar{X})$
TÍPICAS	$Z_x = \frac{X - \bar{X}}{S_x}$	$Z_{y'} = r_{xy} Z_x$	$r_{xy}$ pendiente en típicas	$r_{xy} = \frac{S_{xy}}{S_x S_y}$





COEFICIENTES	SIGNIFICADO	CALCULO
DETERMINACIÓN $r_{xy}^2 = R^2$	Proporción de la varianza de las puntuaciones de los sujetos en el criterio que se puede pronosticar a partir del test	$r_{xy}^2 = \frac{S_{y'}^2}{S_y^2}$
$1 - r_{xy}^2$	Proporción de la varianza de las puntuaciones de los sujetos en el criterio que NO se puede pronosticar a partir del test X	$1 - r_{xy}^2 = \frac{S_{y.x}^2}{S_y^2}$
ALIENACIÓN $CA = K$	<p>Indica la proporción que representa el error típico de estimación respecto a la desviación típica de las puntuaciones en el criterio.</p> <p>Proporción del error típico de estimación respecto a la desviación típica de las puntuaciones en el criterio.</p> <p>Representa la inseguridad, o el azar, que afecta a los Pronostico</p>	$CA = K = \sqrt{1 - r_x^2}$
VALOR PREDICTIVO $CVP$	<p>Capacidad del test para pronosticar el criterio</p> <p>Proporción de seguridad en los pronósticos</p>	$CVP = 1 - \sqrt{1 - r_x^2}$ $CVP = 1 - CA$



	DIRECTAS	DIFERENCIALES	TÍPICAS
VARIANZA TOTAL		$S_y^2 = S_{y'}^2 + S_{y.x}^2$	
ERROR TÍPICO DE MEDIDA	$S_{y.x} = S_y \sqrt{1 - r_x^2}$	$S_{y.x} = S_y \sqrt{1 - r_x^2}$	$S_{zy.zx} = \sqrt{1 - r_x^2}$
ERROR MÁXIMO	$E_{max} = Z_{\alpha/2} S_{y.x}$	$E_{max} = Z_{\alpha/2} S_{y.x}$	$E_{max} = Z_{\alpha/2} S_{zy.zx}$
INTERVALO CONFIDENCIAL	$Y' \pm E_{max}$	$y' \pm E_{max}$	$Z_{y'} \pm E_{max}$



## Ejemplo

	TEST	CRITERIO
1	4	4
2	11	7
3	5	5
4	8	6
5	6	5
6	6	5
7	6	4
8	6	5
9	10	6
10	8	6
11	4	3
12	6	6
13	4	2
14	9	7
15	9	7
16	1	3
17	10	7
18	8	5
19	4	1
20	4	5

En la tabla adjunta aparecen las puntuaciones obtenidas por 20 sujetos en un test X que costa de 25 ítems. Además aparecen las puntuaciones obtenidas por los sujetos en una prueba final de rendimiento que se ha tomado como variable criterio.

- a) Calcular la validez del test para predecir el rendimiento de los alumnos en el examen.
- b) ¿Qué porcentaje de seguridad tendremos en nuestros pronósticos?
- c) ¿Cuál es el error típico de estimación?
- d) ¿Qué puntuación le pronosticaremos en el criterio al sujeto 4, utilizando el método de la regresión?



## Examen (Mayo 2014)

En la tabla siguiente se muestran las puntuaciones obtenidas por 5 sujetos en un test (X) y en un criterio (Y). Con esos datos responder a las preguntas 16- 18 ambas inclusive.

Sujetos	A	B	C	D	E
Test	5	3	6	1	4
Criterio	7	4	5	2	3

Sabiendo que la varianza de los errores de medida del test es igual a 0.59, la varianza del criterio es igual a 2.96 y el coeficiente de validez es 0.76.

16.- El coeficiente de fiabilidad del test es: a) 0.70; b) 0.68; c) 0.80.

17.- El error de estimación cometido con el sujeto C será: a) 1; b) -0.87; c) 5.87.

18.- La proporción de varianza del criterio que no puede explicarse mediante el test es: a) 0.42; b) 0.58; c) 0.24.



## Examen (Mayo 2014)

20.- La desviación típica de los errores de estimación de un test de comprensión verbal es de 0.80 puntos, siendo la media y varianza obtenida por una muestra de sujetos de 8 y 12 respectivamente. La varianza de las puntuaciones de los mismos sujetos en un examen final de Lengua es de 10 puntos y la media 7. Si se quisiera utilizar el test para pronosticar las puntuaciones de los sujetos en el examen de lengua, el porcentaje de seguridad en los pronósticos sería igual al: a) 64%; b) 83%; c) 97%.



## Examen (Junio 2016)

### Con los datos que se presentan a continuación responder a las preguntas 16 a 18

Se quiere averiguar si un test (X) es un buen predictor de la capacidad matemática de los niños de 10 años. Para ello se aplica a una muestra representativa a la que posteriormente se le va a evaluar mediante un examen (Y) en una escala de 0-10. Los resultados obtenidos se muestran a continuación:  $\bar{X} = 10$ ;  $\bar{Y} = 5$ ;  $S_x = 3$ ;  $S_y = 2$  y a partir del test se puede predecir el 49% de la varianza del criterio.

16. La puntuación pronosticada en el examen ( criterio) a un sujeto que en el test obtuvo una puntuación directa de 15 puntos fue: a) 5.33; b) 6.33; c) 7.33
17. El error de estimación de un sujeto que obtuvo en el test una puntuación de 10 puntos y en el examen obtuvo un 6 fue: a) 1; b) 0; c) -1.
18. La proporción de varianza del criterio que no se puede predecir a partir del test es: a) 0,51; b) 0.71; c) 0.29



## Ejemplos

1. Se aplicó un test de fluidez verbal compuesto por 20 elementos a una muestra de 100 sujetos, obteniéndose una desviación típica de 5 y una media de 10 puntos. La correlación entre dicho test y un criterio externo es de 0,80.  
Calcular el coeficiente de alienación, determinación y de valor predictivo. Interpretar los resultados.
2. El 75% de la varianza de un test está asociada a la de su criterio. ¿Cuál será el su coeficiente de alienación?
3. Hemos aplicado un test, con un coeficiente de valor predictivo de 0,34 a un grupo de sujetos. La varianza obtenida es de 144.  
¿Qué puntuación pronosticaríamos en el criterio a un sujeto que ha obtenido en el test una puntuación diferencial de 6 puntos?
4. Se aplicó un test de discriminación visual compuesto de 10 ítems a un grupo de 2.000 sujetos. El coeficiente de fiabilidad del test fue 0,78 y su correlación con un criterio de 0,70. La media del test fue de 6 y la varianza de 4. La media en el criterio fue de 100 y la varianza 25.  
¿Entre qué límites estará la puntuación directa de un sujeto en el criterio, si en el test obtuvo una puntuación de 7 puntos? (NC 99%).



## Ejemplos del libro

**127.** En una muestra de 100 escolares, la desviación típica de los errores de estimación de una prueba de razonamiento matemático es de 0,75 puntos, siendo su media 6 y su varianza 9. Por otro lado, la varianza de las puntuaciones en las calificaciones finales de matemáticas es de 8 puntos, y su media 5. Calcular:

1. El coeficiente de determinación, alienación y el de valor predictivo.
2. Al nivel de confianza del 95% ¿qué calificación se le pronosticará a un sujeto que en la prueba ha obtenido una puntuación de 7 puntos?

**143.** Hemos aplicado un test, con un coeficiente de valor predictivo de 0,34 a un grupo de sujetos. La varianza obtenida es de 144. ¿Qué puntuación pronosticaríamos en el criterio a un sujeto que ha obtenido en el test una puntuación diferencial de 6 puntos?

**148.** Si la varianza de  $X + Y = 61$ , la varianza de  $X - Y = 21$  y la desviación típica de  $X$  es el 80% de la desviación típica de  $Y$ , ¿cuál es el coeficiente de correlación  $r_{xy}$ ?

**191.** Un test tiene un coeficiente de determinación de 0,36. La varianza del test es de 4 puntos y la media 10. La media del criterio es de 5 puntos y la varianza de 25. Un sujeto obtiene en el test una puntuación empírica de 12. Al nivel de confianza del 95%, ¿entre qué valores estará la puntuación de nuestro sujeto en el criterio?