## CONFIABILIDAD Y VALIDEZ EN LA ELABORACIÓN DE INSTRUMENTOS DE MEDICIÓN EN PSICOLOGÍA

¿El instrumento mide lo que dice medir?

¿El instrumento es estable en la medición?

Octubre 2023

Dr. en Ed. Carlos Saúl Juárez Lugo Elaboración de instrumentos Lic. en Psicología V 19 Centro Universitario UAEM Ecatepec Universidad Autónoma del Estado de México Octubre 2023

#### Objetivo de la unidad de aprendizaje.

Diseñar un instrumento de medición psicológica sobre algún constructo, por medio de la identificación de las fases, conceptos, así como procedimientos, con la finalidad de demostrar las cualidades metodológicas requeridas: estandarización, normalización, validez interna y externa y confiabilidad; para su pertinente empleo en las Ciencias Sociales.

## Licenciatura en Psicología Unidad de aprendizaje: Elaboración de instrumentos

- Clave: LPSI14
- Carácter: Obligatorio
- Tipo: Curso taller
- Periodo escolar: Séptimo
- Créditos: 6
- ► Horas teóricas: 2
- Horas prácticas: 2

- ► Área curricular: Metodológica
- Núcleo de formación:

Sustantivo

- Seriación: Ninguna
- UA Consecuente: Ninguna

#### Ubicación de la unidad de aprendizaje en el mapa curricular

MAPA CURRICULAR DE LA LICENCIATURA EN PSICOLOGÍA, 2019								
PERIODO 1	PERIODO 2	PERIODO 3	PERIODO 4	PERIODO 5	PERIODO 6	PERIODO 7	PERIODO 8	PERIODO 9
Bases biológicas del comportamiento I 8	Bases biológicas del 0 comportamiento II 8	Psicología de la 3 1 4 7	Psicopatología del niño y del adolescente 2 4 6	Psicopatología del adulto y del adulto mayor	Diagnóstico 2 2 4 6	Intervención clinica 2 2 2 4 4 6	Intervención clínica 2 en adulto y adulto mayor 4 6	
Procesos del desarrollo y ciclo vital I	Procesos del desarrollo y ciclo vital II 7	Medición y estadística 4 8	Estadística inferencial 4 7	Investigación 0 4 8	Investigación 0 cualitativa 4 8	Elaboración de instrumentos 4 6	Investigación 4 psicológica 6 8	P
Procesos psicológicos básicos 4  8	Procesos psicológicos superiores I 4 8	Procesos psicológicos superiores II 8	Instrumentos psicológicos I 4 6	Instrumentos 2 2 2 4 6	Neuropsicología 2 2 4 6	Integrativa profesional 8		C
Historia de la Psicología  4 0 4 8	Teorías de la personalidad 4 8	Psicología de las organizaciones 4 7	Administración estratégica 4 7	Integración de recursos humanos 2 4 6	Desarrollo de recursos humanos   2   2   4   6	Psicología del 2 2 4 6	Intervención en las 3 organizaciones 4 5	a P
Epistemología 4 0 4 8	Metodología de la ciencia 4 0 4 8	Teorías de la Psicología educativa 3 1 4 7	Proceso de enseñanza-aprendizaje 3 1 4 7	Orientación 2 2 4 6	Evaluación 2 2 4 6	Intervención psicológica inclusiva 6	Intervención y desarrollo educativo 5	o f e
Socialización y vulnerabilidad humana 4	Derechos humanos y bienestar social 4 7	Psicología social 3 1 4 7	Ciudadanía y 3 1 1 4 7	Comunicación y 0 4 8	Proceso grupal 3 1 4 7	Psicología 3 4 5	Procesos psicosociales colectivos  1 3 4 5	s i o n
			Entrevista $\begin{bmatrix} 1\\3\\4\\5 \end{bmatrix}$	Ética en la 2 Psicologia 4 6				a    -  -
	Inglés 5 2 2 4 6	Inglés 6 2 2 4 6	Inglés 7 2 2 4 6	Inglés 8 2 2 4 6				

Unidad de Aprendizaje a la que se destina el material: Elaboración de instrumentos.

#### Secuencia didáctica:

- 1. Evolución histórica de la psicometría.
- 2. Construcción de instrumentos de medición psicológica.
- 3. Evaluación de los instrumentos: confiabilidad y validez

### Guion explicativo

- La confiabilidad y validez en la construcción de un instrumento para medir un constructo en psicología adquiere gran relevancia. El material didáctico que se presenta tiene como objetivo mostrar las ideas más relevantes de las cualidades psicométricas que corresponden a la confiabilidad y validez.
- Se exponen los conceptos y una explicación de lo que evalúa, así como ejemplos de la forma en que se demuestra su reporte en pruebas psicológicas. Además, en el material se mencionan los procedimientos para su cálculo y como se tiene que interpretar el valor numérico que lo acompaña.
- El material está diseñado para acompañar la exposición del docente responsable de la unidad de aprendizaje Elaboración de Instrumentos. Se complementa con otros materiales del mismo autor que se encuentra en el repositorio institucional.



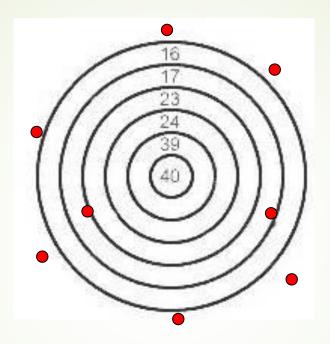
Los principios de validez y confiabilidad en la construcción de instrumentos





Utilizar o Construir un instrumento para la recolección de datos en psicología y su analogía con el deporte de tiro con arco.

Los disparos con las flechas son el equivalente a nuestras preguntas o ítems en un instrumento.



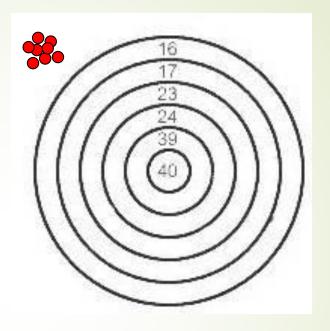
No hay precisión ni tampoco se acierta a donde se debería disparar (al centro del blanco).

Por lo tanto, el tirador no tiene ni confiabilidad ni validez.

(Malos instrumentos)

Los disparos están muy cerca uno de otro, se puede decir que son bastante precisos.

Problema: no está acertando a donde le debería de dar. El tirador tiene problemas de validez, pero no de confiabilidad.



Instrumentos que "miden algo muy bien" pero no lo que pretenden medir.

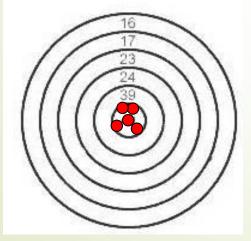


El tirador acierta con precisión exactamente en el blanco. Ha disparado con confiabilidad y validez.

Buenos instrumentos los cuales miden con precisión lo que pretenden medir.

#### Confiabilidad

- Consistencia o estabilidad del instrumento de medición.
- Lø medición es estable



#### Validez

- Mide lo que quiere medir
- ¿El instrumento mide lo que se supone pretende medir?







# A) VALIDEZ

- ¿Estamos midiendo lo que creemos que estamos midiendo?
- ¿El instrumento mide lo que pretende medir?

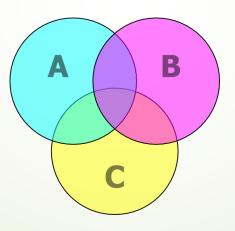
## Principales Tipos

- ► Validez de contenido.
- Validez de criterio.
- Validez de constructo.



#### Validez de contenido

Es la representatividad o la adecuación del muestreo del contenido de un instrumento de medición.



Responde a la pregunta:

¿El **contenido** de esta medida es **representativo** del contenido del **universo** de **contenidos posibles** de la propiedad que se mide?

#### "Rendimiento Aritmético"

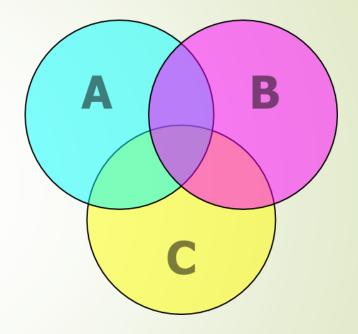
#### Universo

Todos los reactivos posibles utilizando números, operaciones aritméticas y conceptos.

#### Subconjuntos

- A. operaciones aritméticas.
- B. conceptos aritméticos.
- C. manipulaciones numéricas.

- Cualquier muestra de *U* lo suficientemente grande representaría a *A*, *B* y *C* de forma casi igual.
- Esta muestra se materializa en preguntas concretas con un diseño específico.



#### 1.2. FUNDAMENTACIÓN TEÓRICA

Admitida la hipótesis de que los principales procesos cognitivos de procesamiento de información son los de adquisición, codificación o almacenamiento y recuperación, las estrategias cognitivas de aprendizaje o estrategias de procesamiento pueden ser definidas como secuencias integradas de procedimientos o actividades mentales que se activan con el propósito de facilitar

la adquisición, almacenamiento y/o utilización de la información (Nisbett y Shucksmith, 1987).

Tal hipótesis básica es recogida en el modelo de procesamiento de Atkinson y Shiffrin (1968), en la teoría de los niveles de procesamiento de Craik (1979) y Craik y Tulving (1985), en las teorías acerca de

	Factor 1	Subfactor 1 Subfactor 2	Pregunta 01 Pregunta 02 Pregunta 03 Pregunta 01 Pregunta 02
		Subfactor 3	Pregunta 03  Pregunta 01  Pregunta 02  Pregunta 03
	Factor 2	Subfactor 1	Pregunta 01 Pregunta 02 Pregunta 03
Concepto / Constructo		Subfactor 2	Pregunta 01 Pregunta 02 Pregunta 03
		Subfactor 3	Pregunta 01 Pregunta 02 Pregunta 03
	Factor 3	Subfactor 1	Pregunta 01 Pregunta 02 Pregunta 03
		Subfactor 2	Pregunta 01 Pregunta 02 Pregunta 03
		Subfactor 3	Pregunta 01 Pregunta 02 Pregunta 03

Escalas	Subescalas	Estrategias	ÍTEMS
Escala I Estrategias afectivas, de apoyo y control (o	Subescala 1 Estrategias motivacionales (α=.692) (20 ítems)  Subescala 2 Componentes afectivos	Motivación intrínseca ( α=.500)  Motivación extrínseca ( α=.540)  Valor de la tarea ( α=.692)  Atribuciones internas ( α=.537)  Atribuciones externas ( α=.539)  Autoeficacia y expectativas (α=.743)  Concepción de la inteligencia como modificable ( α=.595)  Estado físico y anímico ( α=.735)  Ansiedad ( α=.714)	1, 2, 3 4, 5 6, 7, 8, 9 10, 11, 14 12, 13 15, 16, 17, 18 19, 20 21, 22, 23, 24 25, 26, 27, 28
automanejo) (α=.819) 53 ÍTEMS	(α=.707) (8 items)  Subescala 3  Estrategias metacognitivas ( α=.738) (15 items)	Conocimiento de objetivos y criterios de evaluación (α=.606)  Planificación (α=.738)  Autoevaluación (α=.521)	30, 31 32, 33, 34, 35 29, 36, 39
	Subescala 4 Estrategias de control del contexto, interacción social y manejo de recursos ( \( \alpha =703 \)) (10 ítems)	Control, autorregulación ( α=.660)  Control del contexto ( α=751)  Habilidades de interacción social y aprendizaje con compañeros ( α=.712)	37, 38, 40, 41, 42, 43 44, 45, 46, 47 48, 49, 50, 51, 52, 53

#### Validación de contenido

- Consiste esencialmente en el juicio de expertos.
- El investigador más otros juzgan la representatividad de los reactivos.
- Se puede platear la pregunta:

¿Este reactivo mide la propiedad M?

¿Este reactivo es representativo del universo de contenido de M?

## Se muestra a los expertos

- a) El instrumento o cuestionario en su versión final.
- b) El cuadro de la estructura teórica incluidos los ítems
- c) Guía de evaluación del instrumento.
- d) En ocasiones se trabaja con una "muestra piloto".

#### A-Validez de contenido

Para calcular la validez de contenido mediante el "método de juicio de expertos" se dieron las siguientes instrucciones a 10 jueces, tanto para validar la primera versión de las escalas como para la versión que ahora presentamos:

En el "Departamento de Psicología" de la Universidad de Valladolid, estamos realizando una serie de trabajos sobre Estrategias Cognitivas de Aprendizaje basados, orientados y guiados por modelos de procesamiento de información.

Román, J. M. y Gallego, S. (2001). *Manual ACRA: Escala de Estrategias de Aprendizaje*. Madrid, España: TEA.

Como especialista en Psicología de la Educación, con amplia experiencia en estrategias de aprendizaje, solicitamos tu colaboración. Te pedimos que emitas cuatro calificaciones sobre las Escalas adjuntas, a fin de calcular indicadores subjetivos de validez, a partir de tus puntuaciones, cuantificando de acuerdo con la siguiente escala de respuesta:

*	Nada o en absoluto 0
*	Casi nada 1 ó 2
Þκ	Un poco o algo 3 ó 4
s c	Aceptable 5
겨	Mucho o bastante 6 ó 7
*	Casi absoluta o totalmente 8 ó 9
*	Total y absolutamente 10

(a) ¿En qué medida crees que los 20 elementos de la Escala de Estrategias de La "validez de contenido" (grado en que representa lo que dice representar) calculada mediante el método descrito fue de .87 (análogo de una correlación) para la Escala I, .89 (análogo de una correlación) para la Escala II, .91 (análogo de una correlación) para la Escala III y .88 (análogo de una correlación) para la Escala IV.



#### Validez de criterio

- Se estudia al comparar las puntuaciones de una prueba con una o más variables externas, o criterio, que se sabe o considera que mide el atributo que se estudia.
- Se divide en dos:
  - a) Validez predictiva.
  - b) Validez concurrente.

## A) Validez predictiva

- Involucra el uso de criterios futuros, una prueba nueva vs. una existente.
- Ejemplo. Cuando se predice el fracaso o éxito escolar de los estudiantes a partir de sus medidas de aptitud académica.

¿Qué tan bien predice la prueba EXANI II el promedio final de la licenciatura?

## Tu investigación

- Los instrumentos que utilizarás¿Qué pueden "predecir"?
- Los investigadores que elaboraron el instrumento que han "predicho".
- La "predicción" es una cualidad muy interesante en los instrumentos (regresión lineal).
- Relación estrecha con lo pretende medir.



#### C-Validez predictiva

Se calcularon índices de validez predictiva de las cuatro escalas correlacionando las puntuaciones de cada una con las calificaciones en rendimiento escolar de Ciencias, Historia, Lengua y Matemáticas (según actas):

Escalas	Ciencias	Historia	Lengua	Matemáticas
A.1	0050*	1400**	1 nookk	0.05.0
Adquisición	.0979*	.1482**	.1298**	.0656
Codificación	.1356**	.1458**	.1565**	.0738
Recuperación	.1676**	.2014**	.2126**	.0906*
Apoyo	.1734**	.1991**	.1598**	.0906~

\* p < .05 \*\* p< .01 (dos colas)

Asimismo se realizaron Análisis de Regresión "paso a paso" para hallar las ecuaciones de predicción para cada una de las variables-criterio (rendimiento escolar). Dicho análisis ha seleccionado las variables-predictoras (escalas) que se expresan en las diferentes ecuaciones:

- (a) Ecuación de Predicción para las Ciencias = 3.592 + 0.1733 (Escala de Apoyo al Procesamiento; p< .00001).</p>
- (b) Ecuación de Predicción para la Historia = 3.122 + 0.1237 (Escala de Recuperación; p< .01) + 0.1169 (Escala de Apoyo al Procesamiento; p< .02).</p>

- (c) Ecuación de Predicción para la Lengua = 3.693 + 0.2126 (Escala de Recuperación; p< .00001).</p>
- (d) Ecuación de Predicción para las Matemáticas = 4.840 + 0.1413 (Escala de Recuperación; p< .0003).</p>

Pueden realizarse análisis de regresión más minuciosos calculando las ecuaciones correspondientes pero introduciendo como variables predictoras no las Puntuaciones Directas de cada escala sino las Puntuaciones Directas en cada una de las 32 estrategias.

## B) Validez concurrente

- Mide el criterio casi al mismo tiempo.
- Ejemplo.
- Prueba de inteligencia nueva y se correlaciona con la Stanford Binet, Wechsler.



#### Medición

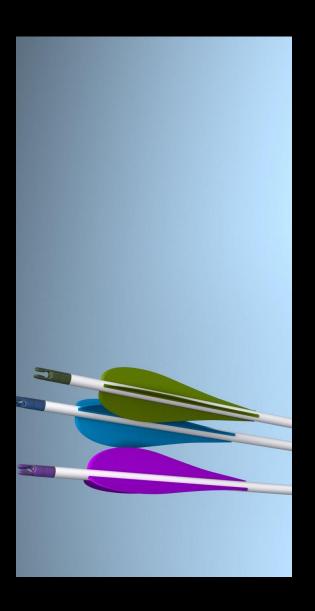
- "Si miden lo mismo los puntajes tendrán correlación entre sí".
- Por medio de pruebas de correlación
- r de Pearson
- Análisis de regresión

$$r = \frac{n(\Sigma XY) - (\Sigma X)(\Sigma Y)}{\sqrt{\left[n(\Sigma \chi^{2}) - (\Sigma \chi)^{2}\right]\left[n(\Sigma y^{2}) - (\Sigma y)^{2}\right]}}$$



### Validez de constructo

- El interés está centrado en las propiedades que se miden, más que en las pruebas utilizadas para lograr la medición.
- Responde a la pregunta:
- ¿Qué propiedades psicológicas o de otro tipo pueden "explicar" la varianza de la prueba



Ejemplo.

Inteligencia.

(Entendido como **constructo**, concepto articulado para tal fin).

¿Qué factores o constructos explican la varianza del desempeño en la prueba?

¿Esta prueba mide habilidad verbal y habilidad de razonamiento abstracto?

- Los investigadores inician con los constructos que tienen relación.
- Intentan validar la teoría que está detrás de la prueba.
- "Inteligencia" vs "Inteligencias Múltiples"

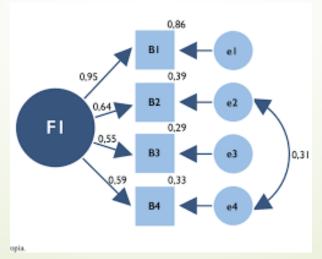


## Evaluación

Por medio del análisis factorial:

Permite identificar si los reactivos diseñados para evaluar el concepto objetivo muestran coherencia de la confirmación de sus

factores.



				Eliminados: 4, 9, 10, 11, 12, 13, 15 y 20				
Repaso	4F03	Factor	Ad	Alfa AF04: 0.726 [12 elementos]		1	2	3
	1	Rmn	18	18. Después de analizar un gráfico o dibujo del texto, dedico algún tiempo a aprenderlo y reproducirlo sin el libro.	AD18	0.703		
	2	Rvoz	16	16. Para comprobar lo que voy aprendiendo de un tema, me pregunto a mí mismo apartado por apartado.	AD16	0.639		
	3	Rvoz	19	19. Hago que me pregunten los subrayados, paráfrasis, esquemas, etc., hechos al estudiar un tema.	AD19	0.618		
	4	Rmn	17	17. Aunque no tenga que presentar examen, suelo pensar y reflexionar sobre lo leído, estudiado, u oído a los profesores.	AD17	0.602		
	5	Rvoz	14	14. Repito la lección como si estuviera explicándosela a un compañero que no la entiende.	AD14	0.580		
Exploración Subrayado	6	SLin	8	8. Empleo los subrayados para facilitar la memorización.	AD8		0.721	
	7	SLin	5	5. En los libros, apuntes u otro material a aprender, subrayo en cada párrafo las palabras, datos o frases que me parecen más importantes.	AD5		0.690	
	8	Slds	7	7. Hago uso de lápices o bolígrafos de distintos colores para favorecer el aprendizaje.	AD7		0.683	
	9	Sids	6	6. Utilizo signos (admiraciones, asteriscos, dibujos), algunos de ellos sólo inteligibles por mí, para resaltar aquellas informaciones de los textos que considero especialmente importantes.	AD6		0.599	
	10	Epg	2	2. Cuando voy a estudiar un material, anoto los puntos importantes que he visto en una primera lectura superficial para obtener más fácilmente una visión de conjunto.				0.743
	11	Exp	1	<ol> <li>Antes de comenzar a estudiar leo el índice, o el resumen, o los apartados, cuadros, gráficos, negritas o cursivas del material a aprender.</li> </ol>	AD1			0.719
	12	Exp	3	3. Al comenzar a estudiar una lección, primero la leo toda por encima.	AD3			0.709



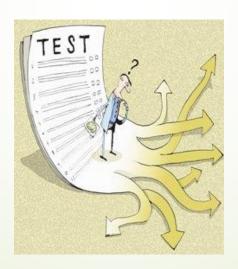
Dr. Carlos Saúl Juárez Lugo 2021

# B) CONFIABILIDAD

La consistencia de un instrumento.

### Confiabilidad

Es un término importante en la elaboración y aplicación de un instrumento de medición en psicología.



### Confiabilidad

Refiere nociones de estabilidad, fiabilidad y predictibilidad.

"Un instrumento confiable es aquel cuya medición es estable, consistente\* y predecible\*".

#### ■ Recuerda

La confiabilidad es un asunto estadístico y de probabilidad

"Acierto / error"



## Definición

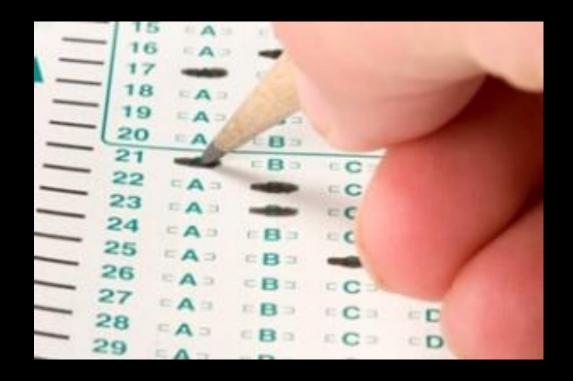
- 1. Se refiere al grado en el que la medición concuerda consigo misma.
- La ausencia relativa de errores de medición en un instrumento de medición.



## **Principios**

- Proceso que indaga que tanto error de medición existe en un instrumento de medición:
  - Mayor confiabilidad menor error de medición.
  - Menor confiabilidad mayor error de medición.





Si una prueba es confiable, solamente se sabe que está midiendo "*algo*" con precisión.

# Métodos para obtener el coeficiente de Confiabilidad

# Coeficiente de correlación de Pearson r

$$r = \frac{n(\Sigma XY) - (\Sigma X)(\Sigma Y)}{\sqrt{\left[n(\Sigma \chi^{2}) - (\Sigma \chi)^{2}\right]\left[n(\Sigma y^{2}) - (\Sigma y)^{2}\right]}}$$

r = coeficiente de correlación de Pearson.

n = número de sujetos.

X = puntuación de la prueba "antes de".

Y = puntuación de la prueba "después de".



Consiste en aplicar el mismo instrumento de medición al mismo grupo de personas, en dos ocasiones diferentes.

# Test / Retest



El lapso de tiempo entre las dos ocasiones depende del tipo y del propósito de las mediciones.



Consiste en desarrollar dos formas equivalentes de la prueba, no es la misma prueba.

Forma equivalentes o paralelas



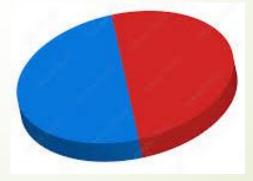
El inconveniente es desarrollar las dos pruebas.

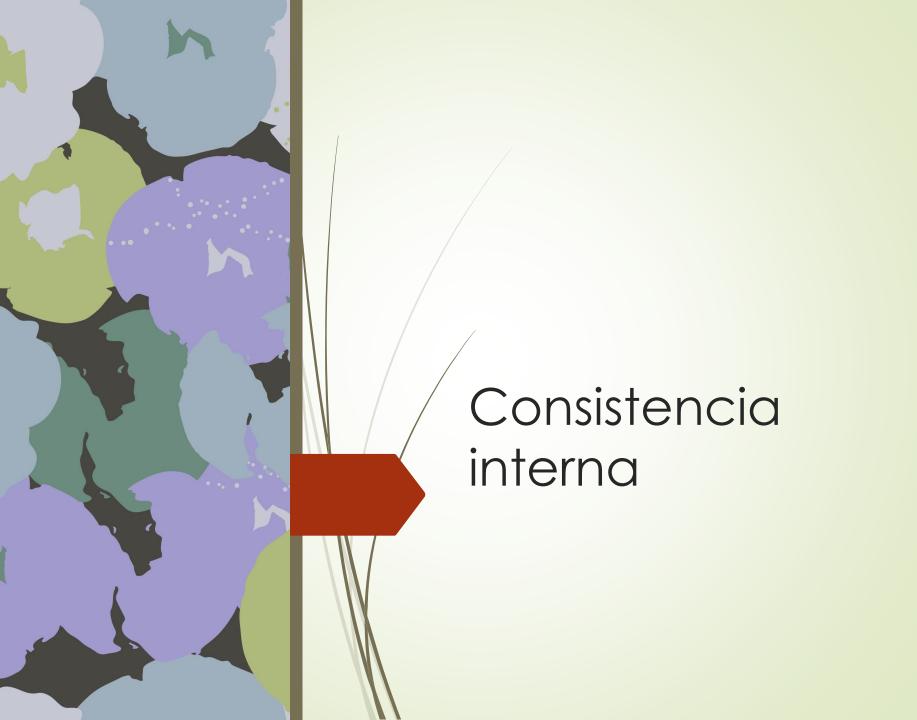


Se recomienda la aplicación en una sola sesión.

La prueba regularmente se divide en dos partes por un lado las preguntas pares y por otro las preguntas nones.

División por mitades





### Consistencia interna

Coeficiente alfa de Cronbach

$$\alpha = \frac{K}{K-1} \left( 1 - \frac{\sum_{i=1}^{K} S_i^2}{S_T^2} \right)$$

Es una forma más general que KR, se utiliza en escalas Likert para determinar la consistencia interna.

### Consistencia interna

- Coeficiente Kuder Richardson
- **KR-20** trabaja por proporciones
- **KR-21** se asume igualdad en kpq (media).
- Son aplicables a instrumentos de medición con un sistema dicotómico o binario de calificación de respuesta.

$$r_{tt} = \frac{n}{n-1} * \frac{Vt - \Sigma pq}{Vt}$$

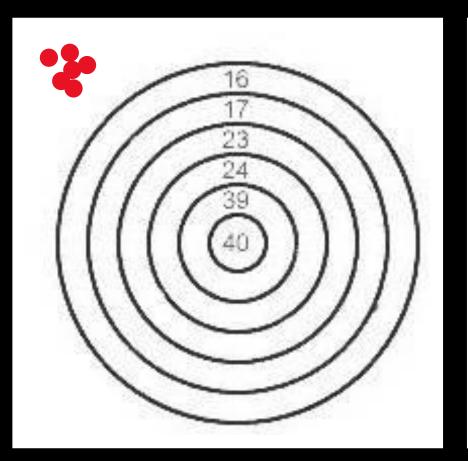
### Interpretación de los valores de alfa de Cronbach

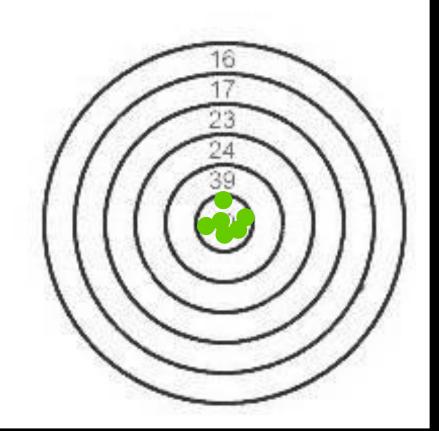
Rangos	Magnitud				
0.81 a 1.00	Muy alta				
0.61 a 0.80	Alta				
0.41 a 0.60	Moderada				
0.21 a 0.40	Ваја				
0.01 a 0.20	Muy baja				
Ruiz-Bolívar, C. (s/f). Confiabilidad.					

Rangos	Magnitud			
0.9	Excelente			
0.8	Bueno			
0.7	aceptable			
0.6	Cuestionable			
0.5	Pobre inaceptable			
Frías-Navarro, D. (2014).				

- Un instrumento puede ser confiable pero no valido.
- Es posible tener confiabilidad sin validez, pero no a la inversa.







Confiable pero no valido

Confiable y valido

### Comentarios finales

- El largo proceso de elaboración de un instrumento, una vez que tenemos el documento en la versión de trabajo y aplicado a una muestra representativa, requiere ser sometido a las primeras pruebas que muestren evidencia científica de que cumple con los criterios de validez y confiabilidad, ahora podemos afirmar que:
  - El instrumento mide lo que dice medir.
  - El instrumento es estable en la medición.

## BLIBLIOGRAFÍA

- Aiken, L. (2003). Test psicológicos y evaluación (11 ed.) México: Pearson Prentice Hall.
- Cohen, R. J. & Swerdlik, M, E. (2001). Confiabilidad. En Introducción a las pruebas y a la medición (pp. 154-183). México: McGraw-Hill.
- Kerlinger, F., y Howard, B. (2002). Investigación del comportamiento. Métodos de investigación en ciencias sociales (4 ed.). México: McGraw-Hill.
- Magnusson, D. (2009). Teoría de los Tests. México: Trillas.
- Nunnaly, J. L. (1991). Teoría Psicométrica. México: Trillas.
- Román, J. M. y Gallego, S. (2001). Manual ACRA: Escala de Estrategias de Aprendizaje. Madrid, España: TEA.

# BIBLIOGRAFÍA COMPLEMENTARIA

- Backhoff, E., Larrazolo, N., y Rosas, M. (2000). Nivel de dificultad y poder de discriminación del Examen de Habilidades y Conocimientos Básicos (EXHCOBA). Revista Electrónica de Investigación Educativa, 2 (1). Recuperado de http://redie.uabc.mx/vol2no1/contenido-backhoff.html.
- Gargallo, B., Suárez, J. y Pérez, C. (2009). El cuestionario CEVEAPEAU. Un instrumento para la evaluación de las estrategias de aprendizaje de los estudiantes universitarios. Revista electrónica de Investigación y Evaluación Educativa, 15(2), 1-31. Recuperado de http://www.uv.es/RELIEVE/v15n2/RELIEVEv15n2\_5.htm