

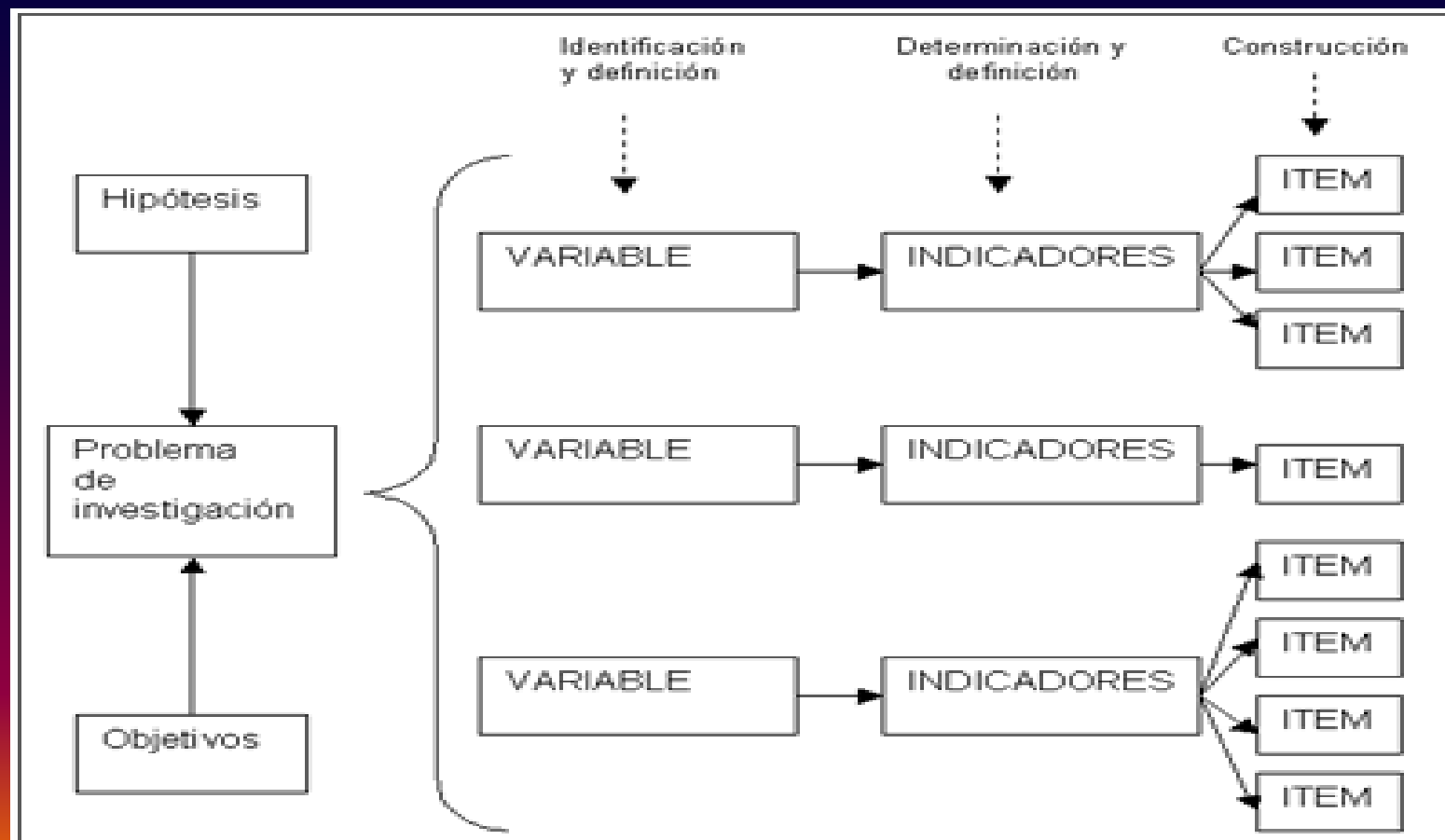
# **INSTRUMENTO PRUEBA PILOTO VALIDACIÓN**

2012

# Preguntas:

- ¿De que depende la selección del instrumento de investigación?
- ¿Cuales son los pasos para construir un instrumento?
- ¿Cuales son las secciones que debe tener el instrumento?
- ¿Todos los instrumentos de recolección de datos de una investigación deben ser validados?
- ¿Que pasos debemos seguir para validar un instrumento que hemos diseñado?
- ¿Todos los pasos para validar un instrumento deben ser aplicados siempre?

# Secuencia lógica para la construcción de un cuestionario



José Casas Jiménez, Jaime García Sánchez, Fernando González Aguilar. Guía técnica para la construcción de cuestionarios. Odiseo. Revista electrónica de pedagogía. México. Año 3, núm. 6. Enero-junio 2006. ISSN 1870-1477

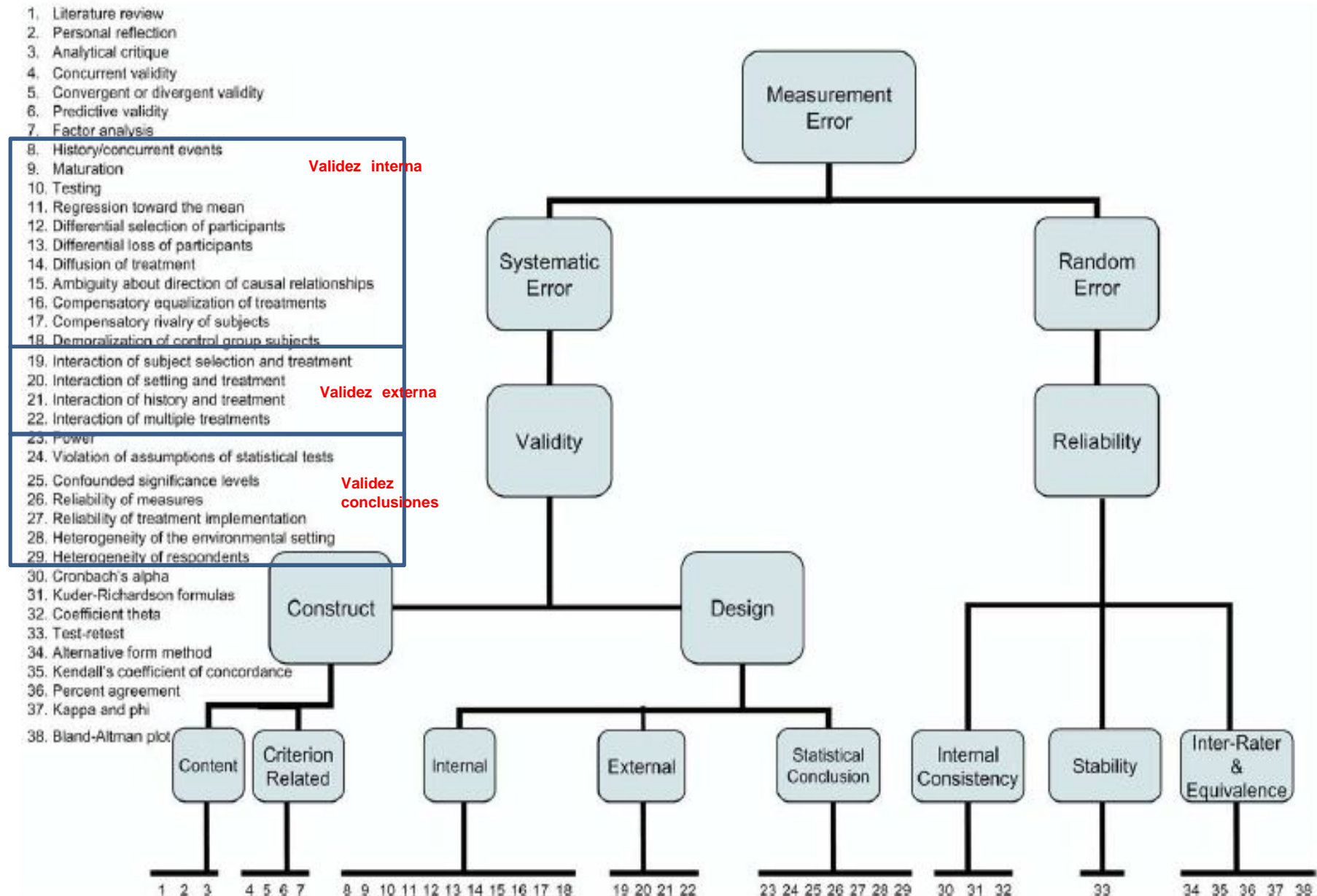


Figure 1. Concept Map

# Prueba piloto

- Es someter a prueba diferentes aspectos de la ejecución de la investigación
- Administrar el instrumento a personas con características de la muestra objeto del estudio
- Disponer en el instrumento de espacio para comentarios
- Realizarlo **antes de aplicar** el instrumento en un estudio
- **Fin:** evaluar aspectos técnicos, mejorar y corregir todas las deficiencias que se encuentren.
- **Número:** Aplicar a **20 o 30 personas** lo mas homogéneas y parecidas posible a los individuos de la muestra (mismo nivel educativo, la misma patología, etc

# Prueba piloto

## Da información acerca de la entrevista:

- Hora más adecuada,
- Tiempo que demora la entrevista.

## Da información acerca del cuestionario:

- *Tipo de pregunta*
- *Orden de las preguntas*
- *Claridad*
- *El diseño del instrumento*

# Tres objetivos principales de la prueba piloto

- Evaluar la idoneidad del cuestionario.
- Calcular la extensión de la encuesta o el tiempo necesario para completarla.
- Determinar la calidad de la tarea del encuestador. (Iraossi 2006, 89).

# Lista de control para el proceso de prueba piloto (Iraossi 2006, 90-92)

- ¿Los encuestados comprenden el objetivo de la encuesta?
- ¿Los encuestados se sienten a gusto al responder las preguntas?
- ¿La redacción de la encuesta es clara?
- ¿La referencia del tiempo es clara para los encuestados?
- ¿Las opciones de respuestas se relacionan con las experiencias de los encuestados en el tema de la pregunta?
- ¿Algunos de los puntos los obligan a pensar demasiado antes de responder? Si es así, ¿cuáles?
- ¿Qué puntos les producen enojo, molestia o confusión?
- ¿Algunas de las preguntas tienen influencia sobre la respuesta? Si es así, ¿cuáles?
- ¿Las respuestas recopiladas reflejan la información que necesita con respecto al objetivo de la encuesta?
- ¿Las respuestas recibidas son lo suficientemente variadas?
- ¿La encuesta es demasiado extensa?
- De acuerdo con sus participantes de prueba, ¿se pasaron por alto otros temas importantes?



# Prueba piloto: aspectos a valorar

| ASPECTO                         | QUE VALORAR   |
|---------------------------------|---|
| ÁREA DE ESTUDIO                 | <ul style="list-style-type: none"> <li>•Seguridad del área</li> <li>•Ambiente (comodidad – privacidad)</li> </ul>   |
| AUTORIZACION                    | <ul style="list-style-type: none"> <li>•Falta de autorización (problemas)</li> </ul>  |
| TIEMPO                          | <ul style="list-style-type: none"> <li>•Tiempo requerido para ejecutar los procesos</li> <li>•Tiempo de traslados</li> </ul>  |
| CAPTACIÓN DE SUJETOS DE ESTUDIO | <ul style="list-style-type: none"> <li>•Acceso a los sujetos del estudio</li> <li>•Porcentaje de rechazos</li> </ul>  |
| INSTRUMENTOS                    | <ul style="list-style-type: none"> <li>•Claridad de las preguntas (lenguaje, redacción, ambigüedades) - Comprensión de preguntas</li> </ul>   |
| PROCEDIMIENTOS                  | <ul style="list-style-type: none"> <li>•Orden de preguntas - Orden de opciones -Tendencias</li> </ul>   |
| COORDINACIÓN Y SUPERVISIÓN      | <ul style="list-style-type: none"> <li>•Claridad de instrucciones - Incongruencias - Secuencia lógica</li> <li>•Claridad de participantes en cuanto a tareas y responsabilidades</li> </ul> |

# FUENTES DE ERROR:

- **Ambiente** de la recolección (frio, calor, privacidad)
- **Situación de los sujetos** (hambre, cansancio, hostilidad, falta de tiempo, inseguridad)
- **Persona que recolecta** la información (genera antagonismo, no conoce los procedimientos del estudio, no interpreta adecuadamente, no percibe bien, siente mucha presión.



Universidad Nacional Mayor de San Marcos  
Facultad de Medicina

# Validación de instrumentos de medición

# VALIDACIÓN DE INSTRUMENTOS

Un instrumento de medición debe ser evaluado en los siguientes aspectos:


- I. Validez**
- II. Confiabilidad**
- III. Adecuada amplitud de rango**
- IV. Sensibilidad al cambio**
- V. Utilidad**

# I. – Validez

- Es el grado en que el instrumento mide aquello que afirma medir
- Cuatro tipos de validez de un instrumento:
  - a. *Validez de apariencia*
  - b. *Validez de constructo*
  - c. *Validez de contenido*
  - d. *Validez de criterio*
- a. *Validez de apariencia* : aceptabilidad social. Jueces que dicen si el instrumento en aspecto mide las cualidades deseadas. Los ítems indican lo que queremos medir (**Valoración por expertos o Jueces**)

# I. - Validez

## *b. Validez de constructo*

- **Constructo** es una idea
- El instrumento refleja la teoría del fenómeno o del concepto que mide.
- Las medidas que resultan de las respuestas del cuestionario pueden ser consideradas y usadas como medición del fenómeno que queremos medir
- **No hay patrón de oro** con el cual comparar el nuevo instrumento  las preguntas de la prueba **son homogéneas** y por lo tanto miden un solo aspecto
- Puede ser calculada por diversos métodos como el **análisis factorial y la matriz multirasgo-multimétodo**

# I. - Validez

## *c. Validez de contenido:*

- Evaluado por un **grupo de expertos o Jueces**
- Instrumento refleja un dominio específico de contenido de lo que se mide. Debe contener a todos los ítems del dominio de contenido de las variables a medir.
- Se debe utilizar un cuestionario con preguntas que posteriormente se llevan a una matriz, a la que se le dará un puntaje.

Prueba de Concordancia entre los Jueces:

$$b = \frac{T_a}{T_a + T_d} \times 100$$

Donde:

$T_a$ : n° total de acuerdos de los jueces

$T_d$ : n° total de desacuerdos de los jueces

$b$ : Grado de concordancia significativa.

|           |            |
|-----------|------------|
| ACEPTABLE | 0.70       |
| BUENO     | 0.70-0.80, |
| EXCELENTE | ≥ 0.90.    |

## ESCALA DE CALIFICACIÓN

Estimado (a):

Teniendo como base los criterios que a continuación se presenta, se le solicita dar su opinión sobre el instrumento de recolección de datos que se adjunta:

Marque con una (X) en SI o NO, en cada criterio según su opinión.

| CRITERIOS  | SI<br>(1) | NO<br>(0) | OBSERVACIÓN |
|--|-----------|-----------|-------------|
| 1. El instrumento recoge información que permite dar respuesta al problema de investigación. |           |           |             |
| 2. El instrumento propuesto responde a los objetivos del estudio.                            |           |           |             |
| 3. La estructura del instrumento es adecuado.<br>( Í tens : 1,2,3,4,5,6,7,8)                 |           |           |             |
| 4. Los ítems del instrumento responde a la operacionalización de la variable.                |           |           |             |
| 5. La secuencia presentada facilita el desarrollo del instrumento.                           |           |           |             |
| 6. Los ítems son claros y entendibles.<br>( Í tens : 1,2,3,4,5,6,7,8)                        |           |           |             |
| 7. El número de ítems es adecuado para su aplicación. ( Í tens : 1,2,3,4,5,6,7,8)            |           |           |             |

### SUGERENCIAS:

(Usado en el INSN)

.....  
.....  
.....  
.....

FIRMA DEL JUEZ EXPERTO (A)



| Criterios    | Jueces   |          |          |          | Valor de P |
|--------------|----------|----------|----------|----------|------------|
|              | J1       | J2       | J3       | J4       |            |
| 1            | 0        | 1        | 1        | 1        | 3          |
| 2            | 0        | 1        | 0        | 1        | 2          |
| 3            | 0        | 1        | 0        | 1        | 2          |
| 4            | 0        | 1        | 0        | 1        | 2          |
| 5            | 0        | 1        | 0        | 1        | 2          |
| 6            | 0        | 1        | 0        | 1        | 2          |
| 7            | 0        | 1        | 0        | 1        | 2          |
| <b>Total</b> | <b>0</b> | <b>7</b> | <b>1</b> | <b>7</b> | <b>15</b>  |

\*: 1 : De acuerdo  
0 : Desacuerdo

Prueba de Concordancia entre los Jueces:

$$b = \frac{T_a}{T_a + T_d} \times 100$$

Procesamiento:

Ta = 15 (N° TOTAL DE ACUERDO DE JUECES)

Td = 13 (N° TOTAL DE DESACUERDO DE JUECES)

b = ( 15 / (15+13)) \*100 = 53.6%. de acuerdo entre los jueces.

ACEPTABLE 0.70

BUENO 0.70-0.80,

EXCELENTE POR ENCIMA DE 0.90.

# Validez de criterio

## d. Validez de criterio:

- Correlación de la escala con otra prueba que mida lo mismo, idealmente con el **patrón de oro** (no siempre existe) que se ha usado rutinariamente y es aceptado.
- Dos tipos: validez concurrente y validez predictiva.

### La validez concurrente

- Comparar las puntuaciones de la prueba con otros indicadores paralelos o criterios externos.
- **Correlación** entre el nuevo instrumento y una escala que ya ha sido probada (patrón de oro) para medir una característica.

Validez Concurrente: X e Y se miden **en el mismo momento**.

Ejm: Realización de mamografía y biopsia (criterio)

# Validez de criterio

## La validez predictiva

- **Grado en que una prueba sirve para predecir comportamientos futuros.**
- **Ecuación de regresión** donde se describe una **línea recta** que se ajusta mejor para la estimación de la característica evaluada a partir del nuevo instrumento.
- **El coeficiente de correlación de Pearson** con valores superiores a 0.8 indica que el nuevo instrumento funciona casi igual al **patrón de oro**
- **El coeficiente de correlación de Pearson** con valores menores de 0.3: el instrumento mide cosas totalmente diferentes al fenómeno que se quiere estudiar.
- Validez de Criterio predictivo:
  - X se mide en el momento 1
  - Y se mide en el momento 2.

Ejemplo:

Test para medir el coeficiente de inteligencia y el rendimiento académico (criterio)

# Confiabilidad

- Capacidad de la escala para proveer los **mismos resultados en mediciones sucesivas**, bajo las mismas características.
- La **constancia o estabilidad de los resultados** que proporciona un instrumento de medida (Bisquerra, 1987)
- **Diferentes mediciones** de un suceso estable en las mismas condiciones deben dar los mismos resultados.
- Llamado también **precisión o reproducibilidad**.
- **Nivel en que los diferentes ítems o preguntas de una escala están relacionados entre sí**. Esta homogeneidad entre los ítems nos indica el **grado de acuerdo entre los mismos**.

# Confiabilidad

Se evalúan tres aspectos: Consistencia interna, Test Retest e Interevaluador

1. **Consistencia interna:** preguntas o ítems que miden un mismo aspecto debe estar fuertemente correlacionados comparados con las preguntas o ítems que miden otro aspecto.

**Correlación muy alto** dice que las dos preguntas están evaluando el mismo aspecto y con uno de los dos es suficiente para medirlo.

Preguntas o ítems miden aspectos diferentes de un mismo fenómeno? = **medidas de consistencia interna** o **medidas de homogeneidad**.

**Alta:** las preguntas están midiendo el mismo aspecto

**Bajo:** están midiendo diferentes fenómenos.

**Prueba de Kuder-Richardson** se conoce como fórmula 20 y se usa para tests con ítems dicotómicos

**Coeficiente Alfa de Cronbach** aplicable a escalas de varios valores posibles

## Ejemplo: Alfa de Cronbach

| Nº Formularios                                 | PREGUNTAS |      |      |      |      |      |      |      | Total   |
|--|-----------|------|------|------|------|------|------|------|---|
|  | p1        | p2   | p3   | p4   | p5   | p6   | p7   | p8   | Sum Fila (t)                                  |
| 1  | 3         | 2    | 3    | 2    | 2    | 2    | 3    | 3    | 20  |
| 2  | 3         | 3    | 2    | 3    | 2    | 2    | 2    | 2    | 19  |
| 3  | 2         | 3    | 2    | 1    | 3    | 2    | 2    | 2    | 17  |
| 4  | 2         | 2    | 2    | 2    | 3    | 2    | 3    | 2    | 18  |
| 5  | 2         | 2    | 3    | 2    | 2    | 3    | 2    | 2    | 18  |
| 6  | 3         | 3    | 3    | 2    | 2    | 2    | 1    | 2    | 18  |
| 7  | 2         | 4    | 3    | 2    | 3    | 1    | 1    | 2    | 18  |
| 8  | 2         | 3    | 2    | 1    | 2    | 1    | 2    | 2    | 15  |
| 9  | 2         | 2    | 2    | 2    | 2    | 2    | 2    | 2    | 16  |
| 10   | 3         | 3    | 3    | 2    | 3    | 2    | 2    | 2    | 20  |
| 11   | 3         | 3    | 3    | 3    | 2    | 2    | 2    | 3    | 21  |
| 12   | 3         | 3    | 3    | 3    | 3    | 3    | 3    | 3    | 24  |
| PROMEDIO columna (i)                           | 2.50      | 2.75 | 2.58 | 2.08 | 2.42 | 2.00 | 2.08 | 2.25 | Varianza total                                |
| DESV EST columna Si                            | 0.52      | 0.62 | 0.51 | 0.67 | 0.51 | 0.60 | 0.67 | 0.45 | columnas                                      |
| VARIANZA por ítem Si <sup>2</sup>              | 0.27      | 0.39 | 0.27 | 0.45 | 0.27 | 0.36 | 0.45 | 0.20 | Varianzas total de ítems St <sup>2</sup> 5.70 |
| SUMA DE VARIANZAS de los ítems Si <sup>2</sup> | 2.65      |      |      |      |      |      |      |      |   |

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

**Donde:**

- $\alpha$  = Coeficiente de Crombach
- $K$  = N° de items utilizados para el cálculo
- $Si^2$  = Varianza de cada item
- $St^2$  = Varianza total de los items

**Reemplazando:**

$$\alpha = (8/8-1) ( 1- ( 2.65/5.70))$$

$$\alpha = (1.1428 * 0.5345)$$

$$\alpha = 0.61$$

# Confiabilidad

## *Alfa de Crombach*

Valores: oscila entre 0.0 y 1.0

0 significa confiabilidad nula

1 representa confiabilidad total

Valores aceptables de alfa

Para propósitos de **investigación**  $\geq 0,7$  (aceptable 0.70, buen índice 0.80 y excelente 0.90).

Fiabilidad **respetable a partir de 0,80.**

Para propósito de **toma de decisiones**  $\geq 0,9$



# Confiabilidad

## 2. **Confiabilidad Test-Retest:** Estabilidad del instrumento a lo largo del tiempo.

Hacer mediciones repetidas a los pacientes. Se recomienda entre 15 días y un mes. Las medidas utilizadas son los **coeficientes de correlación de Pearson y el coeficiente de correlación intraclass** (el más recomendable)

- **Coeficiente de Correlación de Pearson:** no es útil cuando hay más de dos observadores y no mide el grado de acuerdo
- **Coeficiente de correlación intraclass:** mide el **acuerdo (concordancia) entre los diferentes evaluadores**. Una correlación del 70% indicaría una fiabilidad aceptable.

## 3. **Confiabilidad Interevaluador:** Mide que tan **similares** son los **puntajes asignados por diferentes evaluadores** a un mismo fenómeno; para el caso en que los evaluadores entrevistan a las personas.

- Para **valores continuos** (cuantitativos) **coeficiente de correlación intraclass**.
- Para **valores categóricos** (cualitativos) **Índice de Kappa y el porcentaje de acuerdo**.

# Adecuada amplitud de rango

Capacidad del instrumento para **medir todas las características** del suceso de interés.

Un ejemplo de esto es, si se evalúa calidad de vida debe incluir tanto los aspectos físicos como emocionales y sociales, entre otros.

# Sensibilidad al cambio

- Habilidad del instrumento para encontrar las variaciones o cambios que presenta el fenómeno.
- Si mide dolor, los puntajes obtenidos con el instrumento deben disminuir cuando hay mejoría del proceso y viceversa.

Para evaluarla se requiere **aplicar el instrumento de manera repetida al mismo sujeto.**

Se mide por el análisis de covarianza.

# Utilidad

Es la **aplicabilidad** del instrumento de medición en la vida real.

## En resumen con la validez y confiabilidad se pretende:

1. Lo que se valida no es el test, sino las puntuaciones del test, y por lo tanto la pregunta que tratamos de responder es: ¿es válido el uso de las puntuaciones de este test?
2. La validez no se puede resumir en un solo indicador o índice numérico, al igual que ocurre con la confiabilidad (coeficiente de confiabilidad, alfa de Crombach, etc.)
3. La validación es un proceso continuo y dinámico
4. La teoría desempeña un papel muy importante como guía tanto del desarrollo de un test como de un proceso de validación

*Gracias*