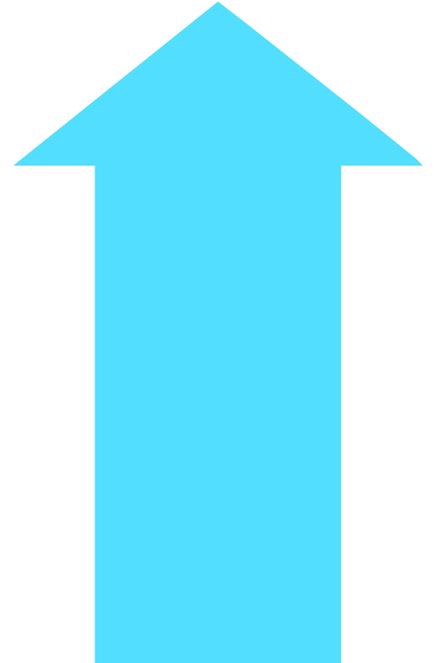
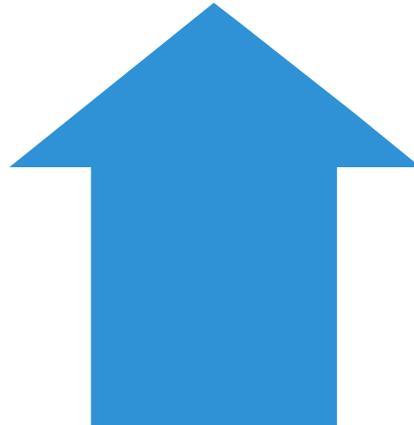
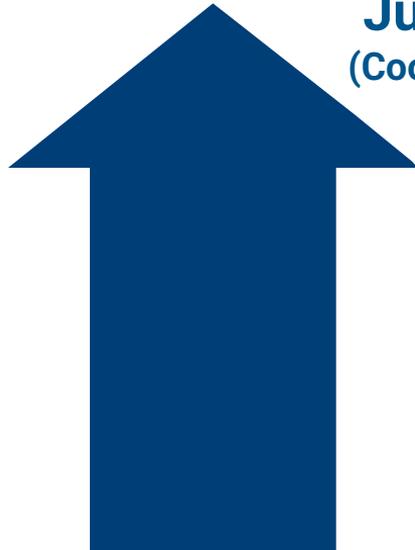
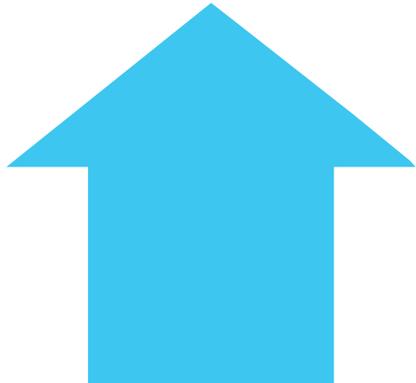


# Seminario 6: fiabilidad y validación de cuestionarios

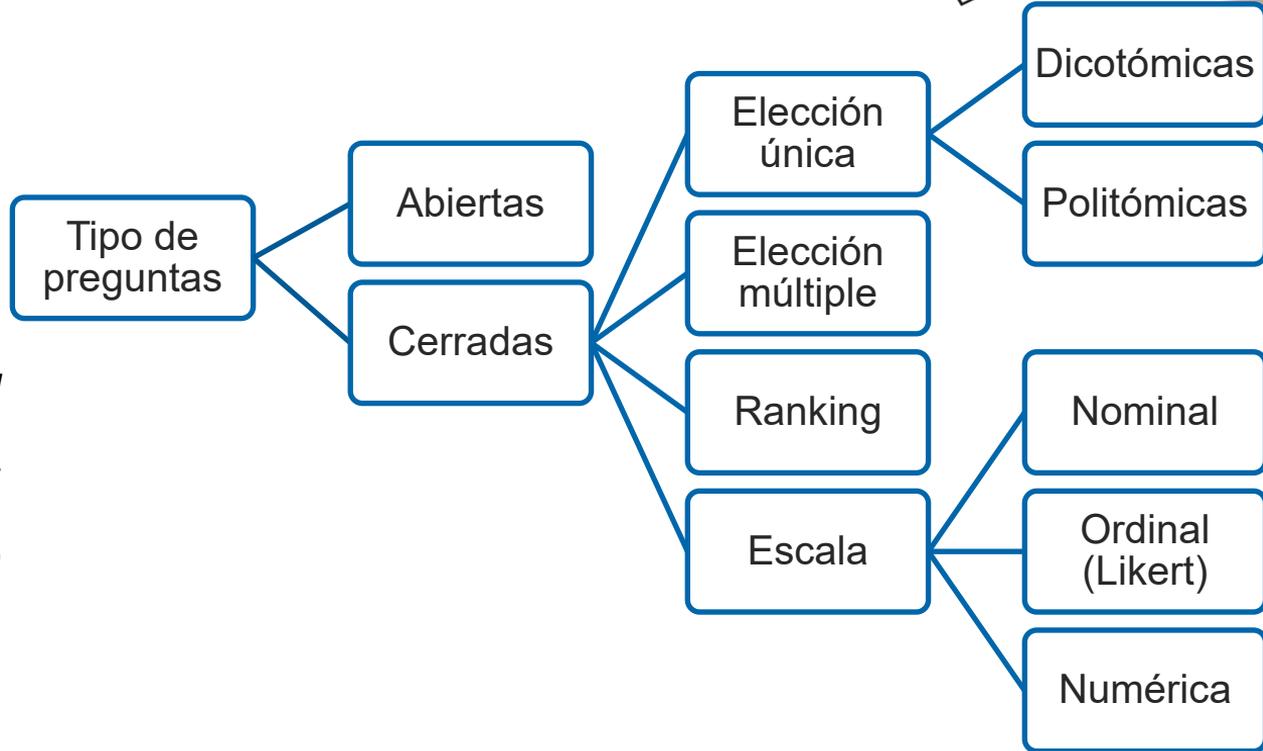


**Julio Ballesta Claver**  
(Coordinador de Investigación)



# 1. Fiabilidad

- Un instrumento de medida (un test o una escala) se considera fiable o confiable si las medidas que se hacen con él *carecen de errores de medida* (son consistentes) esto es, *de si se repite el mismo test con los mismos individuos da el mismo resultado*. Corresponde a la precisión de las medidas.



# 1. Fiabilidad



- Un instrumento de medida (un test o una escala) se considera **fiable** si las medidas que se hacen con él *carecen de errores de medida* (son consistentes) esto es, *de si se repite el mismo test da el mismo resultado*. Corresponde a la precisión de las medidas.
- Más común: la **consistencia interna** de sus ítems (aplicar solo una vez un test).
- Los parámetros que la miden son:
  - **Alfa de Cronbach ( $\alpha$ )**: estima cuánto se relacionan entre sí los ítems de una escala.

Supuestos: unidimensionalidad y varianzas equivalentes.

- Expresión:

$$\alpha = \frac{k}{k-1} \left( 1 - \frac{\sum \sigma_{\text{ítem}}^2}{\sigma_{\text{total}}^2} \right)$$

Limitaciones:

- Sobreestima la fiabilidad si hay muchos ítems redundantes.
- Subestima si los ítems tienen cargas factoriales desiguales.
- No detecta estructuras multidimensionales (otros factores).

Mayor varianza  
“consistente”  
entre ítems =  
mayor fiabilidad

$k$  = número de ítems

$\sigma_{\text{ítem}}^2$  = varianza de cada ítem

$\sigma_{\text{total}}^2$  = varianza de la puntuación total

# 1. Fiabilidad

- Un instrumento de medida (un test o una escala) se considera **fiable** si las medidas que se hacen con él *carecen de errores de medida* (son consistentes) esto es, *de si se repite el mismo test da el mismo resultado*.
- Implica la medida de la **consistencia interna** de sus ítems.
- Los parámetros que lo miden son:
  - **Omega de McDonald ( $\omega$ )**: considera la carga específica de cada ítem sobre el factor común (hace un análisis factorial = multidimensional), sin requerir que todos contribuyan de forma igual (no equivalencia).



$$\omega = \frac{(\sum \lambda_i)^2}{(\sum \lambda_i)^2 + \sum \theta_i}$$

$\lambda_i$  = carga factorial del ítem  $i$

$\theta_i$  = varianza del error del ítem  $i$

# 1. Fiabilidad

Tipo de escala	Nº Ítems	Participantes recomendados
Escala breve (5-10 ítems)	$\geq 5$	$\geq 150$ (alfa), $\geq 250$ (omega)
Escala media (11–20 ítems)	$\geq 10$	$\geq 200$ (alfa), $\geq 300$ (omega)
Escala larga (>20 ítems)	$\geq 20$	$\geq 300$ (mínimo), $\geq 500$ ideal

## Ventajas de omega sobre alfa:

- Se ajusta mejor a escalas con ítems con distinta relevancia.
- Refleja más fielmente la estructura factorial real del cuestionario.
- Aceptado como una medida más moderna y robusta de fiabilidad.



# 1. Fiabilidad

<i>Valor del coeficiente (<math>\alpha</math> y <math>\omega</math>)</i>		<i>Interpretación común</i>
$\geq 0,90$		Excelente (ideal para decisiones individuales)
0,80 – 0,89		Buena (adecuado para decisiones grupales o diagnósticos)
0,70 – 0,79		Aceptable (útil para investigación exploratoria)
0,60 – 0,69		Cuestionable (requiere revisión de ítems)
0,50 – 0,59		Pobre (indica baja consistencia interna)
$< 0,50$		Inaceptable (probable necesidad de rediseñar la escala)

# 1. Fiabilidad con JASP



- El programa **JASP** (Jeffrey's **A**mazing **S**tatistics **P**rogram), en su versión 0.19.3 (última versión), es un programa nombrado en honor a **Harold Jeffreys**, pionero de la inferencia bayesiana. Se trata de un paquete estadístico de código abierto, *fundamentado en R*, desarrollado por un grupo de investigadores de la Universidad de Ámsterdam.

Existe otro como Jamovi que es muy interesante.



[DOWNLOAD](#) | [INFO](#) | [SUPPORT](#) | [TEACHING](#) | [BLOG](#) | [COMMUNITY](#)

¡ Empecemos a utilizar **JASP** ! Descárgalo / Instálalo o Cárgalo

# 1. Fiabilidad con JASP

## 1.1 Poner en castellano



The screenshot shows the JASP software interface with the 'Opciones de la Interfaz de Usuario' (User Interface Options) dialog box open. The dialog box is titled 'Opciones de la Interfaz de Usuario' and contains several sections:

- Tipos de letra**:
  - Interfaz: por defecto: Arial
  - Código R, JAGS o Lavaan: por defecto: Fira Code
  - Resultados y Ayudas: por defecto: Arial
  - Usar el procesamiento de texto de Qt
- Esquemas de colores**:
  - Esquema claro de colores
  - Esquema oscuro de colores
- Idioma preferido**:
  - Escoger idioma: **es - español de España** (highlighted with a red box)
  - [Avúdanos a traducir o mejorar JASP en tu idioma](#)
- Opciones de accesibilidad**:
  - Modo Navegación-ALT
- Opciones diversas**:
  - Zoom (%): - 100 +
  - Velocidad de desplazamiento (pix/s): - 800 +
  - Modo gráfico seguro
  - Deshabilitar las animaciones
  - Utilizar los diálogos de archivos propios

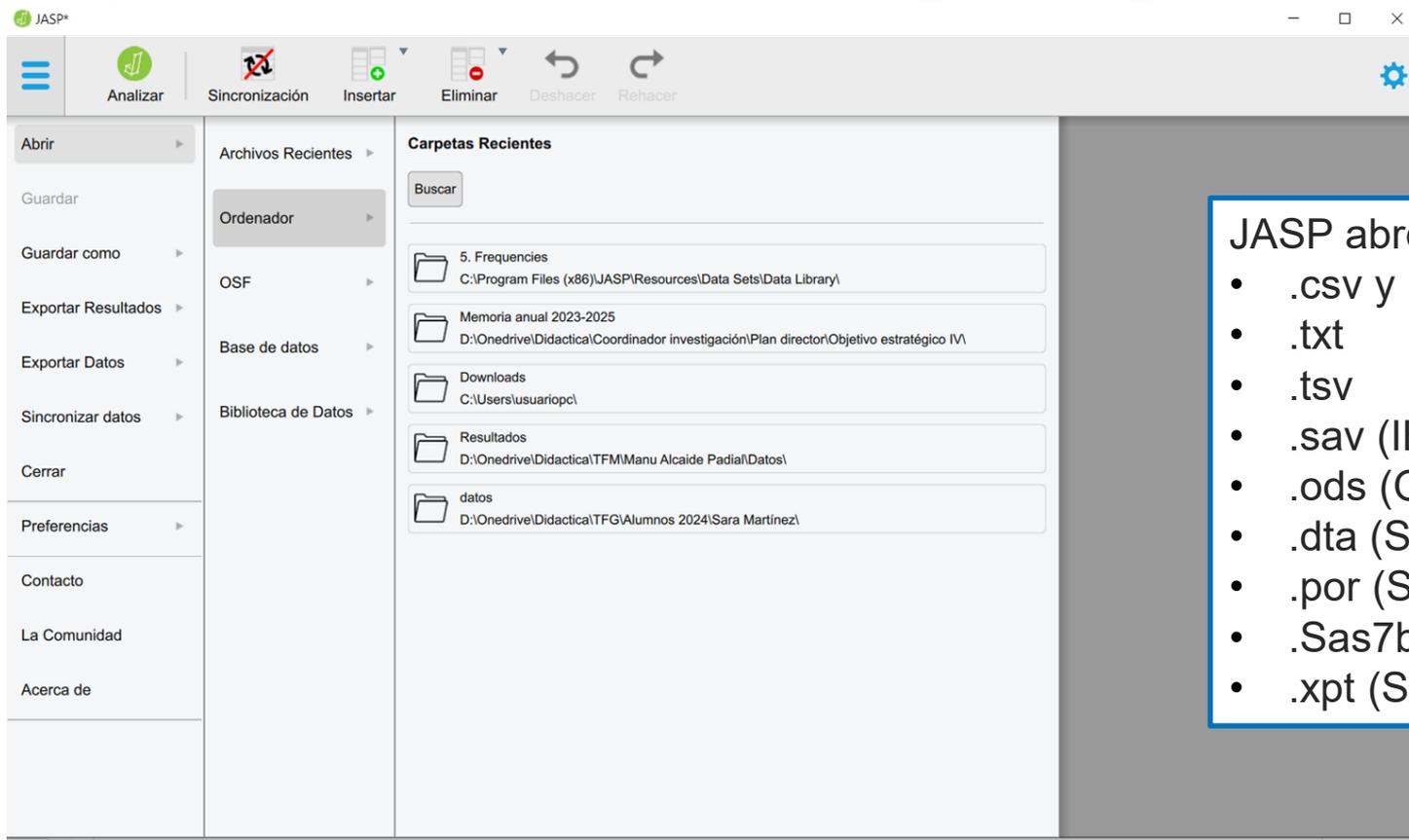
The background of the software interface shows a menu on the left with options like 'Abrir', 'Guardar', 'Exportar Resultados', and 'Preferencias'. The main area displays a landscape graphic with the text 'JASP 0.18.3' and 'usuario.ifestaciones'.

# 1. Fiabilidad con JASP

## 1.2 Importar datos



- Cargar la base de datos que te dará Julio (Base\_Educativa\_Likert.xlsx)



JASP abre los archivos:

- .csv y .xlsx (Excel)
- .txt
- .tsv
- .sav (IBM SPSS)
- .ods (Open office)
- .dta (Stata)
- .por (SPSS ASCII)
- .Sas7bdat /cat (SAS)
- .xpt (SAS)

# 1. Fiabilidad con JASP

## 1.2 Importar datos

Base\_Educativa\_Likert (D:\Onedrive\Didactica\Coordinador investigaci3n\Plan director\Seminarios\Categoría 4\Bibliografía 4\_6\Trabajando con JASP)

	M1	M2	M3	M4	M5	A1	A2	A3	A4	A5	AN1	AN2	AN3	AN4	AN5	E1	E2	E3	E4	E5
1	3	3	3	3	3	4	3	4	3	4	2	3	2	3	3	2	2	2	2	3
2	3	2	3	3	3	3	3	3	3	3	3	2	2	3	3	3	2	2	3	3
3	3	3	3	3	3	2	2	3	2	3	5	3	3	3	4	3	3	4	4	3
4	4	4	4	4	3	3	2	3	3	2	3	3	3	3	3	3	3	3	3	3
5	3	3	3	3	3	2	3	2	2	3	4	3	3	3	4	2	2	2	3	2
6	3	3	3	3	3	3	3	3	2	3	2	2	3	2	3	2	3	3	2	3
7	4	3	3	3	4	3	3	3	3	3	3	3	3	2	2	3	2	2	2	2
8	3	3	3	3	3	3	3	3	3	3	3	2	4	4	4	4	3	4	3	3
9	3	2	3	2	3	2	3	2	3	3	2	2	3	3	2	3	3	3	3	2
10	3	4	4	4	4	4	3	4	3	3	2	2	3	3	3	4	3	3	3	3
11	3	3	2	2	2	3	3	2	3	2	2	3	2	3	3	3	3	3	3	2
12	3	3	3	2	3	3	4	4	4	3	3	3	4	2	4	2	3	3	3	3
13	3	3	3	3	3	4	3	4	3	4	3	3	3	3	3	3	3	2	2	2
14	2	2	3	2	1	3	3	3	3	3	3	2	3	3	3	3	4	4	4	4
15	3	3	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
16	3	3	3	3	3	2	3	2	3	3	3	4	3	3	4	3	3	3	3	3
17	3	2	3	2	3	2	2	2	1	2	3	3	2	3	3	3	4	4	3	2
18	3	3	3	3	3	2	3	3	1	2	3	3	3	2	3	2	2	2	3	3
19	3	3	2	2	2	2	3	3	3	3	4	4	4	4	3	3	3	3	2	3
20	2	2	2	3	3	2	2	3	3	4	3	3	3	3	4	2	3	2	2	2
21	3	3	4	3	4	4	3	4	3	3	3	1	2	2	2	3	3	3	2	3
22	3	3	2	3	3	3	3	3	3	3	3	4	4	3	3	3	3	2	3	2
23	2	4	3	3	3	3	3	2	3	2	2	2	2	2	2	3	4	3	3	3
24	3	3	2	2	2	3	3	3	3	3	2	4	4	3	3	3	3	3	3	3

Base de datos para trabajar:  
**La Competencia psicoeducativa general**, entendida como la capacidad del estudiante para gestionar su motivación, eficacia personal, ansiedad y técnicas de estudio en entornos educativos. Características:

- **300 participantes**
- **20 ítems** tipo Likert (1 a 5)
- **4 dimensiones teóricas:**
  - ✓ **Motivación académica** (M1–M5)
  - ✓ **Autoeficacia** (A1–A5)
  - ✓ **Ansiedad académica** (AN1–AN5)
  - ✓ **Estrategias de estudio** (E1–E5)

# 1. Fiabilidad con JASP

## 1.3 Tipos de datos



Cuadro 1. Tipos de variables según su nivel de medición

Variable	Propiedades	Ejemplos	Representación
<b>Nominal</b>	Mutuamente excluyentes No asignan un orden o jerarquía	Grupo sanguíneo: A, B, y O Sexo: Hombre y Mujer Raza: Blanco, Negro, y Latino	
<b>Ordinal</b>	Establecen un orden, puede ser creciente o decreciente No existe un intervalo número entre las categorías	Escolaridad: primaria, secundaria, bachillerato, licenciatura, maestría y doctorado Riesgo de caídas: alto, medio y bajo	
<b>Intervalo</b>	Establecen un orden determinado por un intervalo numérico. El cero no significa la ausencia de valor	Temperatura: $-10^{\circ}$ , $0^{\circ}$ y $10^{\circ}$	
<b>Razón</b>	Existe un intervalo numérico El cero representa la ausencia de valor, es un cero absoluto.	Numero de hijos: 0, 1, 2... Numero de partos: 0, 1, 2...	



Continua



Fuente: Elaborado a partir de Mendoza y cols, 2011.

# 1.3 Tipos de datos

Nominal



Ordinal



Continua

Base\_Educativa\_Likert (D:\Onedrive\Didactica\Coordinador investigaci3n\Plan director\Seminarior\Categoría 4\Bibliografía 4\_6\Trabajando con JASP)

	M1	M2	M3	M4	M5	A1	A2	A3	A4	A5	AN1	AN2	AN3	AN4	AN5	E1	E2	E3	E4	E5
1	3	3	3	3	3	4	3	4	3	4	2	3	2	3	3	2	2	2	2	3
2	3	2	3	3	3	3	3	3	3	3	3	2	2	3	3	3	2	2	3	3
3	3	3	3	3	3	2	2	3	2	3	5	3	3	3	4	3	3	4	4	3
4	4	4	4	4	3	3	2	3	3	2	3	3	3	3	3	3	3	3	3	3
5	3	3	3	3	3	2	3	2	2	3	4	3	3	3	4	2	2	2	3	2
6	3	3	3	3	3	3	3	3	2	3	2	2	3	2	3	2	3	3	2	3
7	4	3	3	3	4	3	3	3	3	3	3	3	3	2	2	3	2	2	2	2
8	3	3	3	3	3	3	3	3	3	3	3	2	4	4	4	4	3	4	3	3
9	3	2	3	2	3	2	3	2	3	3	3	2	3	3	2	3	3	3	3	2
10	3	4	4	4	4	4	3	4	3	3	2	2	3	3	3	4	3	3	3	3
11	3	3	2	2	2	3	3	2	3	2	2	3	2	3	3	3	3	3	3	2
12	3	3	3	2	3	3	4	4	4	3	3	3	4	2	4	2	3	3	3	3
13	3	3	3	3	3	4	3	4	3	4	3	3	3	3	3	3	3	2	2	2
14	2	2	3	2	1	3	3	3	3	3	3	2	3	3	3	3	4	4	4	4
15	3	3	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
16	3	3	3	3	3	2	3	2	3	3	3	4	3	3	4	3	3	3	3	3
17	3	2	3	2	3	2	2	2	1	2	3	3	2	3	3	3	4	4	3	2
18	3	3	3	3	3	2	3	3	1	2	3	3	3	2	3	2	2	2	3	3
19	3	3	2	2	2	2	3	3	3	3	4	4	4	4	3	3	3	3	2	3
20	2	2	2	3	3	2	2	3	3	4	3	3	3	3	4	2	3	2	2	2
21	3	3	4	3	4	4	3	4	3	3	3	1	2	2	2	3	3	3	2	3
22	3	3	2	3	3	3	3	3	3	3	3	4	4	3	3	3	3	2	3	2
23	2	4	3	3	3	3	3	2	3	2	2	2	2	2	2	3	4	3	3	3
24	3	3	2	2	2	3	3	3	3	3	2	4	4	3	3	3	3	3	3	3

Base de datos para trabajar:  
**La Competencia psicoeducativa general**, entendida como la capacidad del estudiante para gestionar su motivación, eficacia personal, ansiedad y técnicas de estudio en entornos educativos,

- **300 participantes**
- **20 ítems** tipo Likert (1 a 5)
- **4 dimensiones teóricas:**
  - ✓ **Motivación académica** (M1–M5)
  - ✓ **Autoeficacia** (A1–A5)
  - ✓ **Ansiedad académica** (AN1–AN5)
  - ✓ **Estrategias de estudio** (E1–E5)

# 1.3 Tipos de datos



Siempre guarda los cambios

A screenshot of the JASP file menu. The menu is open, showing options: "Nuevo", "Abrir", "Guardar", and "Guardar como". The "Guardar" option is highlighted with a blue arrow. A blue arrow labeled "Asterisco" points to the top right corner of the menu, where a small asterisk icon is located. The window title bar above the menu shows "Base\_Educativa\_Likert (D:". The text "Editar los datos" is visible above the "Nuevo" option.

A screenshot of the JASP main interface. The top toolbar shows icons for "Distribuciones", "Meta análisis", "Fiabilidad", and "Consola de R". The "Fiabilidad" icon is selected. Below the toolbar is a data table with columns labeled "E3", "E4", and "E5". The table contains numerical data. On the right side, there is a sidebar with a list of analysis options. The "Fiabilidad" option is checked. Other options include "Distribuciones", "Contrastes T de Equivalencia", "JAGS", "Aprende Bayes", "Aprender Stats", "Aprendizaje de Máquina", "Meta análisis", "Red", "Potencia", "Análisis Predictivo", "Proceso", "Profeta", "Control de Calidad", and "Contrastes T robustos". A blue arrow points from the text "Siempre guarda los cambios" to the "+" icon in the top right corner of the sidebar.

A screenshot of the sub-menu for the "Fiabilidad" option. The menu is open, showing the "Fiabilidad" icon and the text "Fiabilidad" and "Consola de R". Below this, there are two main categories: "Clásico" and "Bayesiano". Under "Clásico", the option "Fiabilidad Unidimensional" is highlighted with a red rectangle. Other options under "Clásico" include "Correlación Intraclase", "Acuerdo de Calificador", "Gráficos de Bland-Altman", and "Error típico de la medida". Under "Bayesiano", the option "Fiabilidad Unidimensional" is also listed.

# 1. Fiabilidad con JASP



**Fiabilidad Unidimensional**

Variables

M1, M2, M3, M4, M5, A1, A2, A3, A4, A5, AN1, AN2, AN3, AN4, AN5, E1, E2

▼ Análisis

Estadísticos de Escala

Intervalo de confianza 95.0 %

Coeficiente  $\omega$

Coeficiente  $\alpha$

$\lambda^2$  de Guttman

Coeficiente de división en mitades

Correlación media entre elementos

Media  Varianza  DT

de las sumas de las puntuaciones de los participantes

de las medias de las puntuaciones de los participantes

Estadísticos de Elementos Individuales

Intervalo de confianza 95.0 %

Coeficiente  $\omega$  (si el elemento se elimina)

Coeficiente  $\alpha$  (si el elemento se elimina)

$\lambda^2$  de Guttman (al quitar elementos)

Coeficiente de división en mitades (al eliminar el elemento)

Correlación del elemento con el resto

Media  Varianza  DT

✓ Motivación académica (M1–M5)

✓ Estrategias de estudio (E1–E5)

**Resultados**

**Fiabilidad Unidimensional**

Estadísticas de confiabilidad de la escala frecuente

Coeficiente	Estimar	Error típico	IC del 95%	
			Lower	Upper
Coeficiente $\omega$	0.210	0.045	0.122	0.297
Coeficiente $\alpha$	0.605	0.058	0.491	0.719

Nota. Los siguientes ítems se correlacionan negativamente con la escala: M1, M2, M3, M4, M5, E1, E2, E3, E4, E5.

Estadísticas de confiabilidad de la escala frecuente

Coeficiente	Estimar	Error típico	IC del 95%	
			Lower	Upper
Coeficiente $\omega$	0.407	0.043	0.323	0.492
Coeficiente $\alpha$	0.622	0.051	0.521	0.723

Nota. Los siguientes ítems se correlacionan negativamente con la escala: A1, A2, A3, A4, A5.

Estadísticas de confiabilidad de ítems individuales frecuentes

Ítem	Coeficiente $\omega$ (si ítem eliminado)			Coeficiente $\alpha$ (si ítem eliminado)		
	Estimar	IC inferior al 95%	IC superior 95%	Estimar	IC inferior al 95%	IC superior 95%
A1	0.091	0.020	0.162	0.593	0.476	0.711
A2	0.083	0.015	0.152	0.593	0.473	0.713
A3	0.342					
A4	0.342					
A5	0.086					
AN1	0.244					
AN2	0.245					
AN3	0.212					
AN4	0.249					
AN5	0.248					
E1	0.250					
E2	0.271					
E3	0.245					
E4	0.255					
E5	0.264					
M1	0.223					
M2	0.206					
M3	0.211					
M4	0.212					
M5	0.215					

Elementos a Escala Normal

M1, M2, M3, M4, M5, A1, A2, A3, A4, A5, AN1, AN2, AN3, AN4, AN5

Elementos a Escala Invertida

E1, E2, E3, E4, E5

Nota. Los siguientes ítems se correlacionan negativamente con la escala: M1, M2, M3, M4, M5, E1, E2, E3, E4, E5.

Invertirlos empeora el resultado (parece que no están inversos)

Constructo multidimensional y heterogéneo

# 1. Fiabilidad con JASP



## ▼ Fiabilidad Unidimensional

Agrupémoslos por las dimensiones ya establecidas



Al hacerlos por separado,  $\alpha$  y  $\omega$  se equiparan con buen resultado

### ▼ Fiabilidad Unidimensional M

Variables

- M1
- M2
- M3
- M4
- M5

▼ Análisis

### Fiabilidad Unidimensional M

#### Estadísticas de confiabilidad de la escala frecuente

Coeficiente	Estimar	Error típico	IC del 95%	
			Lower	Upper
Coefficient $\omega$	0.764	0.022	0.722	0.806
Coefficient $\alpha$	0.762	0.033	0.697	0.828

Fiabilidad de consistencia interna

#### Estadísticas de confiabilidad de ítems individuales frecuentes

Ítem	Coeficiente $\omega$ (si ítem eliminado)			Coeficiente $\alpha$ (si ítem eliminado)		
	Estimar	IC inferior al 95%	IC superior 95%	Estimar	IC inferior al 95%	IC superior 95%
M1	0.731	0.681	0.781	0.730	0.645	0.814
M2	0.731	0.681	0.781	0.728	0.643	0.814
M3	0.724	0.673	0.775	0.722	0.648	0.796
M4	0.717	0.665	0.769	0.713	0.634	0.792
M5	0.705	0.651	0.760	0.702	0.610	0.794

...cuánto se relacionan entre sí los ítems de una escala...

# 1. Fiabilidad con JASP

## 1.4 Fiabilidad de acuerdo entre jueces



Criterio de valoración	Descripción	Tipo de escala	Justificación
<b>Claridad</b>	¿El ítem está redactado de forma comprensible, sin ambigüedad?	Ordinal (Likert 1–5)	Evalúa la calidad lingüística y comunicativa del ítem.
<b>Pertinencia</b>	¿El ítem es adecuado y relevante para la dimensión que pretende medir?	Ordinal (Likert 1–5)	Evalúa el alineamiento con el constructo.
<b>Relevancia</b>	¿El ítem aporta información significativa?	Ordinal (Likert 1–5)	Evalúa si el ítem es necesario y no redundante.
<b>Ubicación dimensional</b>	¿El ítem corresponde a la dimensión teórica asignada?	Nominal (asignación de dimensión: A, B, C...)	Se usa para analizar consistencia con el modelo teórico.

*Lo ideal es emplear de entre 5 a 10 jueces*

# 1. Fiabilidad con JASP

## 1.4 Fiabilidad de acuerdo entre jueces



MÉTODO	TIPO DE DATO	Nº JUECES	INTERPRETA EL ACUERDO SOBRE...	CRITERIOS DE INTERPRETACIÓN COMÚN
<b>Kappa de Cohen</b>	Nominal 	2	Clasificación entre 2 jueces	< 0,40 = bajo 0,41–0,60 = moderado > 0,60 = bueno
<b>Kappa de Fleiss</b>	Nominal 	$\geq 3$	Clasificación en más de 2 jueces	< 0,40 = pobre 0,41–0,60 = moderado > 0,60 = sustancial
<b>Alfa de Krippendorff</b>	Nominal / Ordinal / Intervalar 	$\geq 2$	Cualquier tipo de escala y dato	0,67 aceptable > 0,80 ideal
<b>Correlación Intraclase (ICC)</b>	Ordinal / Cuantitativo	$\geq 2$	Puntuaciones continuas/ordinales	0,75 = buena > 0,90 = excelente
<b>Kendall's W</b>	<b>Ordinal</b> 	$\geq 2$	Rango o jerarquía de los ítems	0,50–0,70 = moderado 0,70 = fuerte



# 1. Fiabilidad con JASP

## 1.4 Fiabilidad de acuerdo entre jueces



Criterio de valoración	Descripción	Tipo de escala	Métodos recomendados	Nº Jueces	Criterios de interpretación
<b>Claridad</b>	¿El ítem está redactado de forma comprensible, sin ambigüedad?	Ordinal (Likert 1–5)	ICC, Kendall's W, Krippendorff $\alpha$ (ordinal)	$\geq 2$	ICC > 0,75 Kendall's W > 0,70 $\alpha$ > 0,67
<b>Pertinencia</b>	¿El ítem es adecuado y relevante para la dimensión que pretende medir?	Ordinal	 ICC, Kendall's W, Krippendorff $\alpha$ (ordinal)	$\geq 2$	(idem)
<b>Relevancia</b>	¿El ítem aporta información significativa?	Ordinal	ICC, Kendall's W, Krippendorff $\alpha$ (ordinal)	$\geq 2$	(idem)
<b>Ubicación dimensional</b>	¿El ítem corresponde a la dimensión teórica asignada?	Nominal (A, B, C...)	 Kappa de Cohen (pares), Kappa de Fleiss ( $\geq 3$ ), Krippendorff $\alpha$ (nominal)	$\geq 2$ o $\geq 3$	Kappa > 0,60 $\alpha$ > 0,67

# 1. Fiabilidad con JASP

## 1.4 Fiabilidad de acuerdo entre jueces



Criterio de valoración	Descripción	Tipo de escala	Métodos recomendados	Nº Jueces	Criterios de interpretación
<b>Ubicación dimensional</b>	¿El ítem corresponde a la dimensión teórica asignada?	Nominal (A, B, C...)	Kappa de Cohen (pares), Kappa de Fleiss ( $\geq 3$ ), Krippendorff $\alpha$ (nominal)	$\geq 2$ o $\geq 3$	Kappa > 0,60 $\alpha$ > 0,67

- Cargar la base de datos que te dará Julio (Base\_validación\_interjueces\_CON\_RUBRICA.xlsx)

*La base de datos tiene tres jueces: Juez 1, Juez 2 y Juez 3*



# 1. Fiabilidad con JASP

## 1.4 Fiabilidad de acuerdo entre jueces



Base\_validación\_interjueces\_CON\_RUBRICA (D:\Onedrive\Didactica\Coordinador investigaci3n\Plan director\Seminarior\Categoría 4\Bibliografía 4\_6\Trabajando con JASP)

Editar los datos | Descriptivos | Contrastes T | ANOVA | Modelos Mixtos | Regresi3n | Frecuencias | Factor | Estadísticos Circulares | Distribuciones | Meta análisis | **Fiabilidad**

Ítem	Juez1_Claridad	Juez2_Claridad	Juez3_Claridad	Juez1_Pertinencia	Juez2_Pertinencia	Juez3_Pertinencia	Juez1_Relevancia	Juez2_Relevancia	Juez3_Relevancia	Categoría_Juez1	Categoría_Juez2	Categoría_Juez3
1	3	3	2	1	2	1	4	4	4	C	C	C
2	3	3	3	4	4	4	2	2	2	A	A	A
3	3	3	3	1	1	1	3	3	3	C	C	C
4	2	1	2	4	3	4				C	C	C
5	4	4	4	4	4	4				A	A	A
6	4	4	4	4	4	4				A	A	A
7	4	4	4	3	3	3				C	C	A
8	4	3	4	3	3	3				B	A	A
9	3	3	3	3	3	3				C	C	C
10	2	2	3	1	1	2				C	B	C
11	2	2	2	4	4	4				C	C	C
12	3	3	3	3	4	3				C	C	C
13	2	2	2	3	3	3				A	B	A
14	3	3	3	1	1	1				C	C	C
15	4	4	4	3	3	3				B	B	B
16	3	3	3	1	1	1				A	A	A
17	4	4	4	2	1	2				B	B	B
18	4	4	4	3	3	3	3	4	3	B	B	B
19	1	1	1	2	2	2	2	2	2	B	B	B
20	3	3	2	1	1	1	4	3	4	B	A	B

Fiabilidad

- Fiabilidad
- Clásico
- Fiabilidad Unidimensional
- Correlaci3n Intraclase
- Acuerdo de Calificador
- Gráficos de Bland-Altman
- Error típico de la medida



### Acuerdo de Calificador

Item

- Juez1\_Claridad
- Juez2\_Claridad
- Juez3\_Claridad
- Juez1\_Pertinencia
- Juez2\_Pertinencia
- Juez3\_Pertinencia
- Juez1\_Relevancia
- Juez2\_Relevancia
- Juez3\_Relevancia

Variables

- Categoría\_Juez1
- Categoría\_Juez2
- Categoría\_Juez3

Kappa de Cohen
  Intervalo de Confianza 95.0 %

No ponderado
  Ponderado

Ponderaciones cuadráticas
  Ponderaciones lineales

Kappa de Fleiss

Alfa de Krippendorff

Estructura de los Datos
   
 Calificadores en columnas
   
 Calificadores en filas

Método: Nominal

Nº de muestras de bootstrap para IC 1000
   
 Establecer semilla 1234

## Resultados

### Acuerdo de Calificador

#### Kappa de Cohen

Calificaciones	Kappa	ET	IC del 95%	
			Inferior	Superior
Average kappa	0.747			
Categoría_Juez1 - Categoría_Juez2	0.695	0.133	0.435	0.954
Categoría_Juez1 - Categoría_Juez3	0.848	0.100	0.651	1.000
Categoría_Juez2 - Categoría_Juez3	0.697	0.134	0.435	0.959

Nota. 20 sujetos/elementos y 3 calificadores/medidas. Los intervalos de confianza son asintóticos.

ET: Error Típico

#### Kappa de Fleiss

Calificaciones	Kappa de Fleiss	ET	IC del 95%	
			Inferior	Superior
Overall	0.746	0.092	0.566	0.926
A	0.683	0.129	0.430	0.936
B	0.672	0.129	0.440	0.925
C	0.863	0.129	0.610	1.110

Nota. 20 sujetos/elementos y 3 calificadores/medidas. Los intervalos de confianza son asintóticos.

#### Krippendorff's alpha

Método	Krippendorff's alpha	ET	IC del 95%	
			Inferior	Superior
Nominal	0.748	0.099	0.542	0.900

Nota. 20 sujetos/elementos y 3 calificadores/medidas.

### Criterios de interpretación

Kappa > 0,60  
 $\alpha > 0,667$

Tipo de datos evaluados	Kappa de Cohen sugerido
Categoricos (A, B, C)	No ponderado
Ordinales (Likert 1-5)	Ponderado (mejor aún si es cuadrático)

# 1. Fiabilidad con JASP



## 1.4 Fiabilidad de acuerdo entre jueces

1. El **Kappa de Fleiss**, que evalúa el acuerdo global entre los tres jueces, fue de  $\kappa = 0,746$ , con valores elevados en todas las categorías evaluadas (A:  $\kappa = 0,683$ ; B:  $\kappa = 0,672$ ; C:  $\kappa = 0,863$ ), siendo la categoría C la más acordada. Se observa una consistencia general en la clasificación.

2. El **Kappa de Cohen** mostró un acuerdo medio-alto entre los diferentes pares de jueces: Juez<sub>1</sub>–Juez<sub>2</sub> ( $\kappa = 0,695$ ), Juez<sub>1</sub>–Juez<sub>3</sub> ( $\kappa = 0,848$ ) y Juez<sub>2</sub>–Juez<sub>3</sub> ( $\kappa = 0,697$ ), con un promedio global de  $\kappa = 0,747$ . Según los criterios de Landis y Koch (1977), estos valores indican un acuerdo **sustancial a casi perfecto**.

3. Complementariamente, el **alfa de Krippendorff**, índice robusto frente a la distribución de las categorías, arrojó un valor de  $\alpha = 0,748$  (IC 95%: 0,517–0,935), considerado adecuado para su uso en investigación social aplicada.

En conjunto, estos resultados indican que existe un **alto grado de acuerdo interjueces**, lo que aporta **evidencia sólida de validez de contenido** para los ítems analizados.

# 1. Fiabilidad con JASP

## 1.4 Fiabilidad de acuerdo entre jueces



Criterio de valoración	Descripción	Tipo de escala	Métodos recomendados	Nº Jueces	Criterios de interpretación
<b>Claridad</b>	¿El ítem está redactado de forma comprensible, sin ambigüedad?	Ordinal (Likert 1–5)	ICC, Kendall's W, Krippendorff $\alpha$ (ordinal)	$\geq 2$	ICC > 0,75 Kendall's W > 0,70 $\alpha > 0,67$
<b>Pertinencia</b>	¿El ítem es adecuado y relevante para la dimensión que pretende medir?	Ordinal	 ICC, Kendall's W, Krippendorff $\alpha$ (ordinal)	$\geq 2$	(idem)
<b>Relevancia</b>	¿El ítem aporta información significativa?	Ordinal	ICC, Kendall's W, Krippendorff $\alpha$ (ordinal)	$\geq 2$	(idem)

$\alpha$  de Krippendorff  
(ordinales)



## CRITERIOS DE INTERPRETACIÓN COMÚN

$$\alpha > 0,67$$

$\alpha$  de Krippendorff (ordinales)

### Acuerdo de Calificador nominales

### Acuerdo de Calificador ordinales Claridad

Item

- Juez1\_Pertinencia
- Juez2\_Pertinencia
- Juez3\_Pertinencia
- Juez1\_Relevancia
- Juez2\_Relevancia
- Juez3\_Relevancia
- Categoría\_Juez1
- Categoría\_Juez2
- Categoría\_Juez3

Variables

- Juez1\_Claridad
- Juez2\_Claridad
- Juez3\_Claridad

Kappa de Cohen

No ponderado

Ponderado

Ponderaciones cuadráticas

Ponderaciones lineales

Kappa de Fleiss

Alfa de Krippendorff

Estructura de los Datos

Calificadores en columnas

Calificadores en filas

Método Ordinal

Nº de muestras de bootstrap para IC: 1000

Establecer semilla 1234

### Acuerdo de Calificador ordinales Claridad

Krippendorff's alpha

Método	Krippendorff's alpha	ET	IC del 95%	
			Inferior	Superior
Ordinal	0.906	0.059	0.748	0.983

Nota. 20 sujetos/elementos y 3 calificadores/medidas.

### Acuerdo de Calificador ordinales Pertinencia

Krippendorff's alpha

Método	Krippendorff's alpha	ET	IC del 95%	
			Inferior	Superior
Ordinal	0.935	0.040	0.834	0.982

Nota. 20 sujetos/elementos y 3 calificadores/medidas.

### Acuerdo de Calificador ordinales Relevancia

Krippendorff's alpha

Método	Krippendorff's alpha	ET	IC del 95%	
			Inferior	Superior
Ordinal	0.946	0.036	0.851	0.987

Nota. 20 sujetos/elementos y 3 calificadores/medidas.

### Acuerdo de Calificador ordinales Pertinencia

### Acuerdo de Calificador ordinales Relevancia



- ▶ Acuerdo de Calificador nominales
- ▶ Acuerdo de Calificador ordinales Claridad
- ▶ Acuerdo de Calificador ordinales Pertinencia
- ▶ Acuerdo de Calificador ordinales Relevancia

### Correlaci3n Intraclase Claridad

Item

- Juez1\_Pertinencia
- Juez2\_Pertinencia
- Juez3\_Pertinencia
- Juez1\_Relevancia
- Juez2\_Relevancia
- Juez3\_Relevancia
- Categoría\_Juez1
- Categoría\_Juez2
- Categoría\_Juez3

Variables

- Juez1\_Claridad
- Juez2\_Claridad
- Juez3\_Claridad

Cada sujeto se califica mediante...

- un evaluador distinto (seleccionado aleatoriamente)
- el mismo conjunto de evaluadores/contrastes seleccionados aleatoriamente
- el mismo conjunto fijo de evaluadores/pruebas
- Las calificaciones son promedios

Intervalo de Confianza 95.0 %

- ▶ Correlaci3n Intraclase Pertinencia
- ▶ Correlaci3n Intraclase Relevancia

### Correlaci3n Intraclase Claridad

Correlaci3n intraclase

Tipo	Estimaci3n por Punto	IC inferior al 95%	IC superior 95%
ICC3,1	0.901	0.805	0.956

*Nota.* 20 sujetos y 3 calificadores/medidas. Tipo de CCI seg3n la referencia de Shrout & Fleiss (1979).

### Correlaci3n Intraclase Pertinencia

Correlaci3n intraclase

Tipo	Estimaci3n por Punto	IC inferior al 95%	IC superior 95%
ICC3,1	0.940	0.878	0.974

*Nota.* 20 sujetos y 3 calificadores/medidas. Tipo de CCI seg3n la referencia de Shrout & Fleiss (1979).

### Correlaci3n Intraclase Relevancia

Correlaci3n intraclase

Tipo	Estimaci3n por Punto	IC inferior al 95%	IC superior 95%
ICC3,1	0.940	0.878	0.974

*Nota.* 20 sujetos y 3 calificadores/medidas. Tipo de CCI seg3n la referencia de Shrout & Fleiss (1979).



**Clásico**

Fiabilidad Unidimensional

Correlaci3n Intraclase

Acuerdo de Calificador

Gráficos de Bland-Altman

Error t3pico de la medida

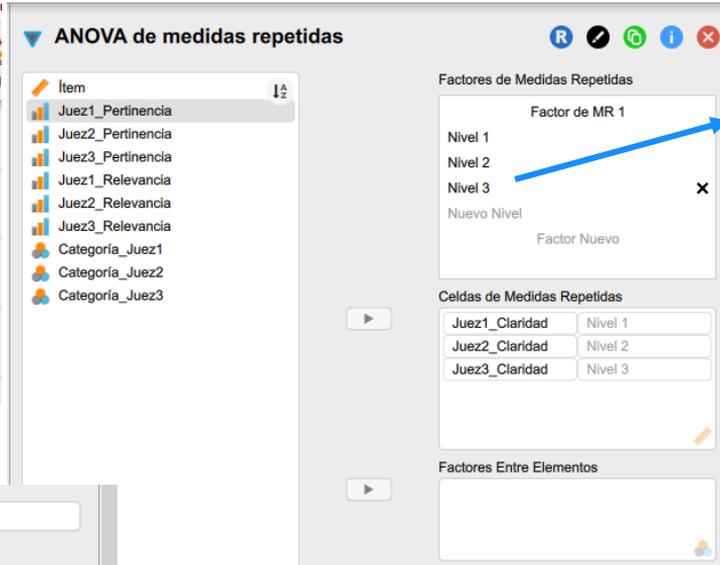
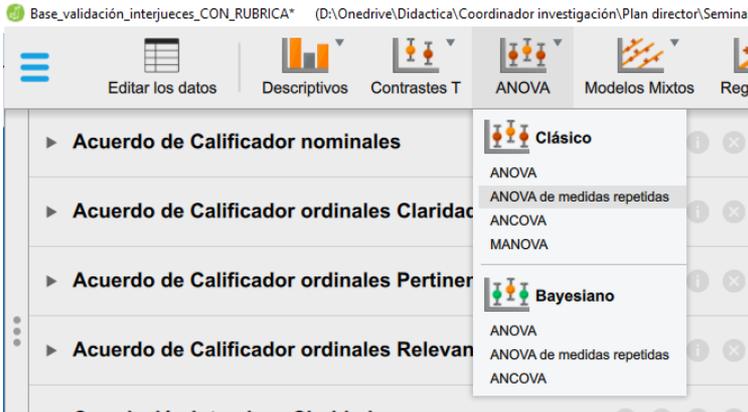
**CRITERIOS DE INTERPRETACI3N COM3N**

ICC > 0,75

**Correlaci3n Intraclase (ICC)**

# 1. Fiabilidad con JASP

## 1.4 Fiabilidad de acuerdo entre jueces

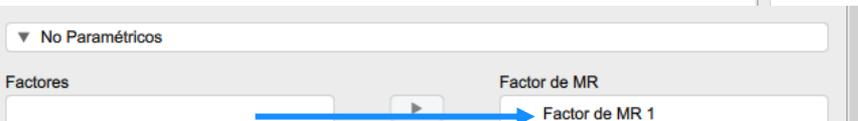


Crear un nivel (nivel 3)

**CRITERIOS DE INTERPRETACIÓN COMÚN**

Kendall's  $W > 0,70$

W de Kendall



El coeficiente  $W$  de Kendall mide el acuerdo en el **orden de los ítems**, no la coincidencia en puntuaciones.

Tus jueces **coinciden en puntuaciones**, pero **no** en la jerarquía relativa entre ítems, por eso  $W$  es bajo.

### No Paramétricos

Contraste de Friedman

Factor	$X^2_F$	gl	p	W de Kendall
Factor de MR 1	1.200	2	0.549	0.030



# 1. Fiabilidad con JASP

## 1.4 Fiabilidad de acuerdo entre jueces



### ¿QUÉ QUIERES EVALUAR?

USA...

¿Coinciden los jueces al asignar **categorías**?

Kappa de Cohen / Fleiss /  
Krippendorff (nominal)

¿Coinciden los jueces en las **puntuaciones numéricas**?

ICC / Krippendorff (ordinal)

¿Coinciden los jueces en el **orden** (jerarquía de ítems)?

W de Kendall

¿Quieres un índice robusto para cualquier tipo de dato?

Alfa de Krippendorff

Ítem	Juez1	Juez2	Juez3
A	1	2	3
B	2	3	4
C	3	4	5

**$\alpha$  de Krippendorff o ICC**  
(ordinales) se centran en  
los valores, no en el  
orden

Ítem	Juez1	Juez2	Juez3
A	2	2	2
B	4	1	3
C	1	4	1

- **W Kendall alto** (orden jerárquico).
- **$\alpha$  de Krippendorff bajo** (porque hay distancia entre valores).

- **$\alpha$  de Krippendorff alto** (porque los números están cerca).
- **W Kendall bajo** (no hay acuerdo en el orden de los ítems).

***¡Gracias por  
vuestra  
asistencia!  
(Parte 1)***

