



GRADO EN PSICOLOGÍA

PSICOMETRÍA

FIABILIDAD Y VALIDEZ: RENDIMIENTO TÍPICO I

MARZO 2021

EDUARDO GARCÍA GARZON

 **Universidad
Camilo José Cela**

1: INTRODUCCIÓN A LA BASE DE DATOS

1.1. Introducción

En esta práctica vamos a analizar una base de datos que contiene las respuestas de 250 sujetos a una serie de test. En esta práctica vamos a centrarnos en el análisis de una serie de ítems de rendimiento típico, analizando la fiabilidad de las puntuaciones del test y las evidencias de validez pertinentes. En esta práctica no analizaremos *las evidencias de validez referidas a la estructura interna*. Las mismas vienen referidas en la tercera guía de Jamovi de la asignatura.

Los test definidos en esta base de datos son los siguientes:

1. **Brief Multidimensional Life Satisfaction Scale:** Escala Breve de Satisfacción Vital Multidimensional. Escala de 12 ítems que mide doce áreas diferentes de satisfacción vital. Cada ítem está medido en escala 1-7. Se tienen las respuestas de los sujetos en el momento de recogida de datos original y a los seis meses después de la recogida de los datos.
2. **The Energy/Activation Scale:** Escala de Energía/Activación física breve de 5 ítem. Esta escala evalúa el actual estado físico subjetivo de cada uno de los sujetos. Cada ítem está medido en escala 1-7.
3. **The Scale of Positive and Negative Experiences (SPANE), subescala positiva:** 6 ítems que miden experiencias positivas del test SPANE. Cada ítem está medido en escala 1-7.
4. **Puntuaciones en Autonomía y Sentimientos Negativos.** Puntuaciones sumadas provenientes de las escalas ACS-30 y la subescala negativa del SPANE.

1.2. Plan de Análisis

Importante

Para realizar esta práctica correctamente es necesario actualizar Jamovi a la versión 1.8.0

La primera parte de la práctica vamos a analizar diferentes aspectos de la **fiabilidad** de las puntuaciones del test BMLSS:

1. Estadísticos de fiabilidad para las puntuaciones totales en la escala medida con los estadísticos α de Cronbach y el estadístico ω de McDonald.
2. Análisis de dificultad y varianza para cada uno de los ítems de la escala.
3. Análisis de resultados del α de Cronbach y el estadístico ω de McDonald si se considera eliminar cada uno de los ítems.
4. Análisis de fiabilidad por otros métodos: **método de las dos mitades y método de fiabilidad test-retest.**

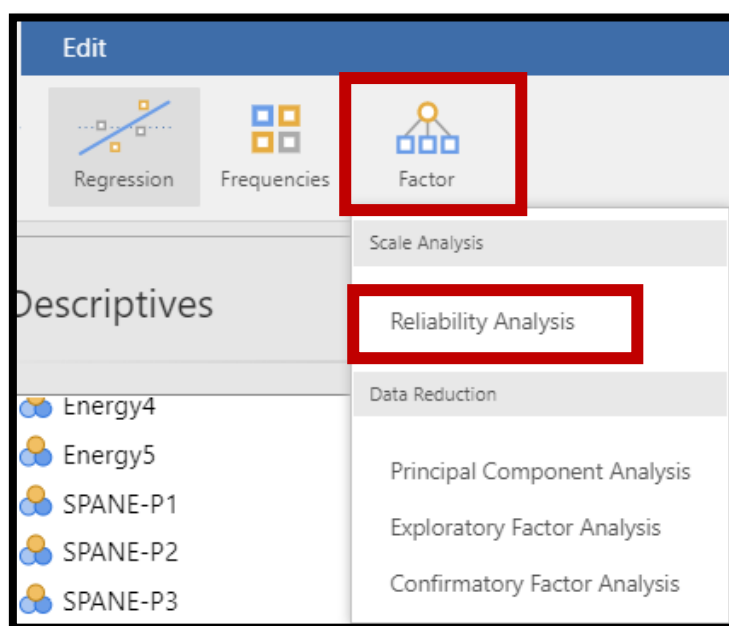
La segunda parte de la práctica vamos a centrarnos en analizar evidencias de **validez** en relación de las puntuaciones de un test y un criterio.

5. Vamos a calcular el coeficiente de validez entre las puntuaciones del test y las puntuaciones de un criterio (puntuaciones en la subescala de experiencias positivas del test SPANE)
6. Vamos a analizar **la validez incremental** de las puntuaciones del BMLSSS cuando intentamos predecir la subescala positiva del test SPANE una vez considerada la influencia de las puntuaciones del test de Energía.
7. Por último, vamos a analizar **la validez congruente y discriminante** cuando consideramos las relaciones con las puntuaciones de Autonomía y la subescala de sentimientos negativos del SPANE.

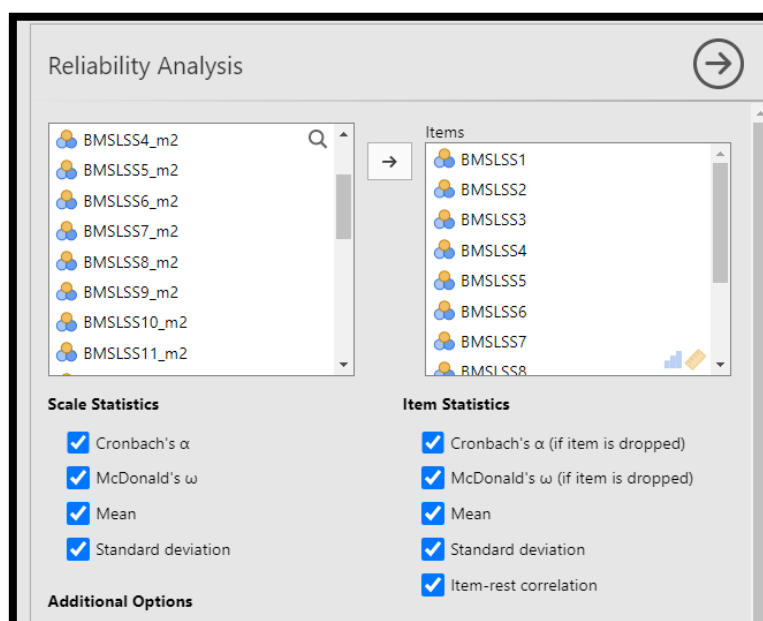
1: Fiabilidad

1.1. Análisis de la fiabilidad de las puntuaciones de la escala

La mayoría de los análisis que vamos a utilizar para analizar la fiabilidad de la escala van a encontrarse dentro del menú **factor** de Jamovi. Dentro del menú **factor**, vamos a centrarnos en el submenú **Reliability Analysis**.



Dentro del menú de análisis tenemos varias opciones que nos van a ser de utilidad para responder a las preguntas que nos interesa responder. Primero vamos a analizar **los estadísticos descriptivos y fiabilidad** de las puntuaciones de la escala, añadiendo las variables BMLSS1-BMLSS12 al cuadro de *Items*.

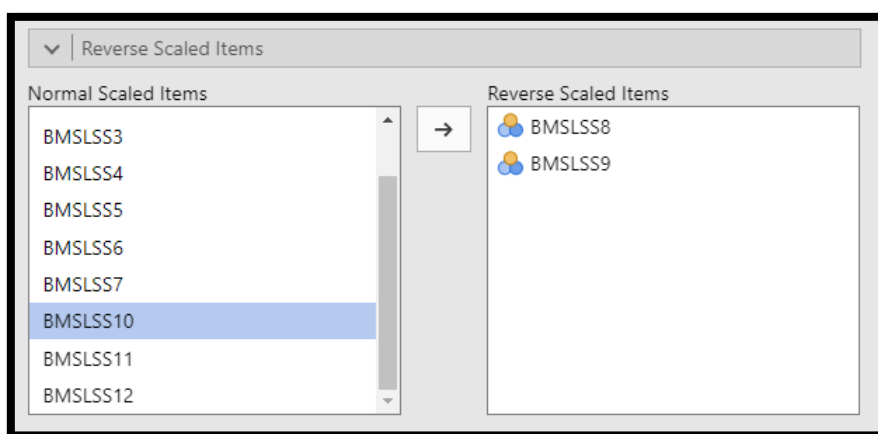


Una vez hayamos añadido las variables correspondientes, podemos solicitar los estadísticos de media, desviación típica, el α de Cronbach y el estadístico ω de McDonald para las puntuaciones de la escala.

Reliability Analysis				
Scale Reliability Statistics				
	mean	sd	Cronbach's α	McDonald's ω
scale	4.412	0.692	0.510	0.620
Note. items 'BMSLSS8' and 'BMSLSS9' correlate negatively with the total scale and probably should be reversed				
[3]				

En este caso hemos observado una media de 4.645, una desviación típica de 0.925. Esto nos indica que la mayoría de las personas tienen un bienestar por encima de la media esperada de la escala (3,5). Por otra parte, los estadísticos α de Cronbach y el estadístico ω de McDonald indican que, según los criterios de la EFPA, no hemos observado una fiabilidad adecuada en la escala.

En este caso particular, Jamovi nos avisa que existen dos ítems que podemos considerar cuantificar de una manera alternativa. En este caso particular, estos ítems son dos ítems que tenemos que **recodificar**, ya que son ítems inversos que nos hemos cuantificado correctamente previamente. Para recodificar estos ítems, podemos utilizar la función **crear variable calculada** y utilizar la fórmula matemática $(K+1) - X_i$ donde K es el número de categorías de la variable. Sin embargo, podemos recodificar estos ítems utilizando el menú de fiabilidad de Jamovi.



Tras recodificar estos dos ítems, podemos observar el cálculo correcto de los estadísticos para las puntuaciones totales de la escala:

Reliability Analysis				
Scale Reliability Statistics				
	mean	sd	Cronbach's α	McDonald's ω
scale	4.646	0.925	0.765	0.774

[3]

En este caso, observamos como cambian la media y la desviación típica de las puntuaciones. Además, observamos que tras recodificar estos dos ítems, los estadísticos α de Cronbach y el estadístico ω de McDonald indican que, según los criterios de la EFPA, hemos observado una fiabilidad adecuada en la escala. Hay que tener en cuenta que esta fiabilidad tiene que interpretarse en consideración con la longitud de la escala (que en este caso es de 12 ítems).

1.2. Análisis de la fiabilidad de las puntuaciones de cada ítem

Si hemos seleccionado el cálculo de los **estadísticos de ítems (Item Statistics)**, obtendremos la media, desviación típica, correlación ítem-retest (correlación ítem-test corregida) y los índices de fiabilidad si se elimina cada uno de los ítems.

	mean	sd	item-rest correlation	if item dropped	
				Cronbach's α	McDonald's ω
BMSLSS1	4.660	1.797	0.338	0.756	0.767
BMSLSS2	4.432	1.967	0.378	0.752	0.765
BMSLSS3	4.884	1.693	0.430	0.746	0.758
BMSLSS4	4.304	1.784	0.429	0.746	0.757
BMSLSS5	5.152	1.752	0.409	0.748	0.757
BMSLSS6	4.408	1.395	0.469	0.744	0.747
BMSLSS7	4.660	1.775	0.336	0.756	0.767
BMSLSS8 *	5.500	2.013	0.378	0.753	0.765
BMSLSS9 *	3.900	1.731	0.392	0.750	0.761
BMSLSS10	4.304	1.791	0.404	0.749	0.759
BMSLSS11	5.140	1.761	0.420	0.747	0.756
BMSLSS12	4.404	1.448	0.478	0.743	0.746

* reverse scaled item

En este caso, observamos que el ítem con el que las participantes mostraron un mayor nivel de acuerdo fue el ítem BMLSS8, y el ítem con un menor nivel de acuerdo fue el ítem BMLSS9. El ítem con una mayor varianza fue el ítem BMLSS9, y el ítem con una menor variabilidad fue el ítem BMLSS6. Hay que recordar que Jamovi nos devuelve la desviación típica de las puntuaciones de cada ítem, no la varianza.

Por otra parte, vemos que el ítem con un mejor índice de discriminación fue el ítem BMLSS12, y el ítem con un peor índice de discriminación el ítem BMLSS7. Ninguno de los ítems mostraba una **correlación ítem-test corregida menor a 0.30**, por lo que no consideramos eliminar ninguno de ellos. Respecto a los índices de fiabilidad si eliminamos los ítems, **no observamos que eliminar ninguno de ellos mejoraría la fiabilidad de la escala** (ni medida por los estadísticos α de Cronbach y el estadístico ω de McDonald), por lo que consideraríamos mantener la escala con sus 12 ítems originales. En ambos casos, el ítem BMLSS12 parece ser el ítem de mayor calidad.

1.3. Análisis de la fiabilidad por el método de dos mitades

Para poder calcular la fiabilidad por el **método de dos mitades**, es necesario que obtengamos las puntuaciones correspondientes a cada mitad del test. En este caso, vamos a seguir un esquema de división por ítems pares e impares. Para ello, calcularemos dos nuevas variables, sumando las puntuaciones de ítems pares e impares. Antes de ello, es necesario que **recodifiquemos** las puntuaciones de los ítems 8 y 9, ya que eran ítems inversos. Para ello, podemos aplicar utilizar la opción de **Calcular Variable** y aplicar la fórmula antes indicada (donde 8 indica el número de categorías totales +1):

COMPUTED VARIABLE

BMSLSS8_rev

Description

Formula

$f_x = 8 - \text{BMSLSS8}$

↑

>

Después de aplicar la misma corrección al ítem 9, podemos calcular la suma de puntuaciones totales para los ítems pares e impares.

COMPUTED VARIABLE

pares

Description

Formula

$f_x = \text{BMSLSS2} + \text{BMSLSS4} + \text{BMSLSS6} + \text{BMSLSS8_rev} + \text{BMSLSS10} + \text{BMSLSS12}$

↑

>

COMPUTED VARIABLE

impares

Description

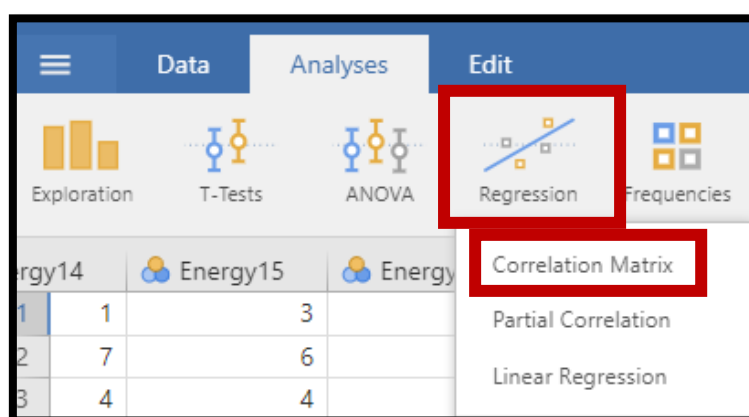
Formula

$$f_x = \text{BMSLSS1} + \text{BMSLSS3} + \text{BMSLSS5} + \text{BMSLSS7} + \text{BMSLSS9_rev} + \text{BMSLSS11}$$

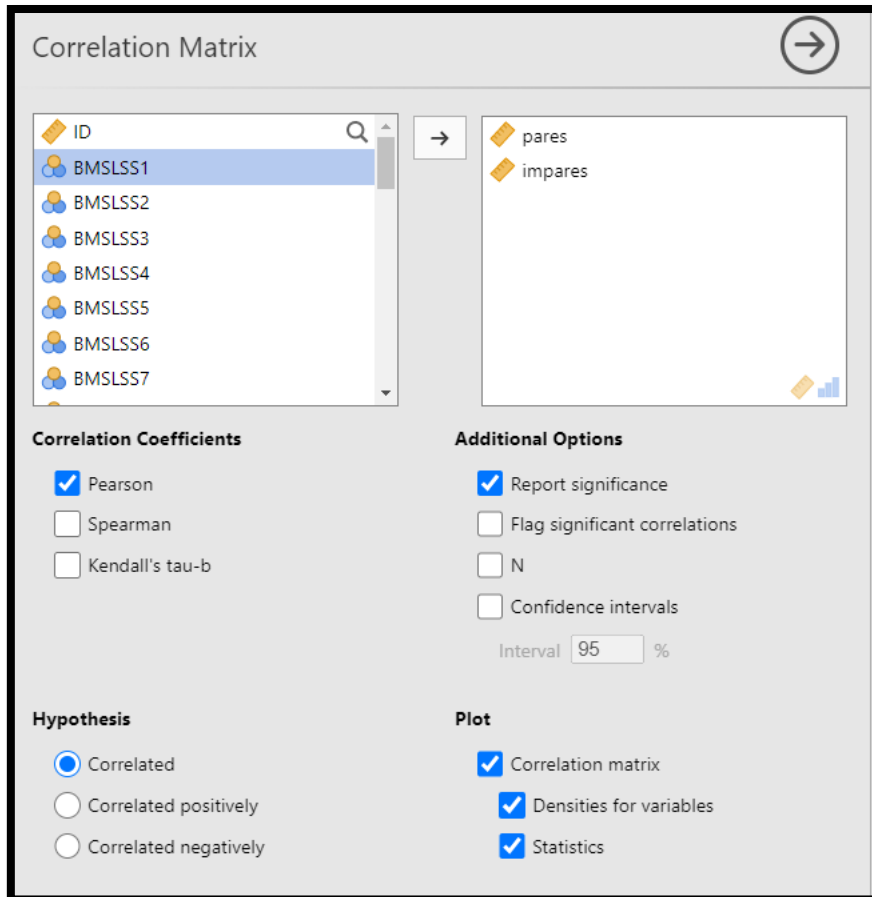
↑

>

Una vez calculadas estas puntuaciones, podemos obtener la **fiabilidad** de cada una de las partes del test calculando la correlación entre estas puntuaciones. Para ello, vamos a utilizar el menú **Regression >> Correlation matrix**.



Dentro de este menú, vamos a introducir las puntuaciones calculadas dentro del menú de variables, dejando el resto de las opciones que nos marca por defecto Jamovi. Es útil obtener en este caso una visualización adecuada de la relación entre las puntuaciones de los ítems pares e impares.



Correlation Matrix

Variables: BMSLSS1, BMSLSS2, BMSLSS3, BMSLSS4, BMSLSS5, BMSLSS6, BMSLSS7

Correlation Coefficients:

- ☒ Pearson
- ☐ Spearman
- ☐ Kendall's tau-b

Additional Options:

- ☒ Report significance
- ☐ Flag significant correlations
- ☐ N
- ☐ Confidence intervals
- Interval: 95 %

Hypothesis:

- ☒ Correlated
- ☐ Correlated positively
- ☐ Correlated negatively

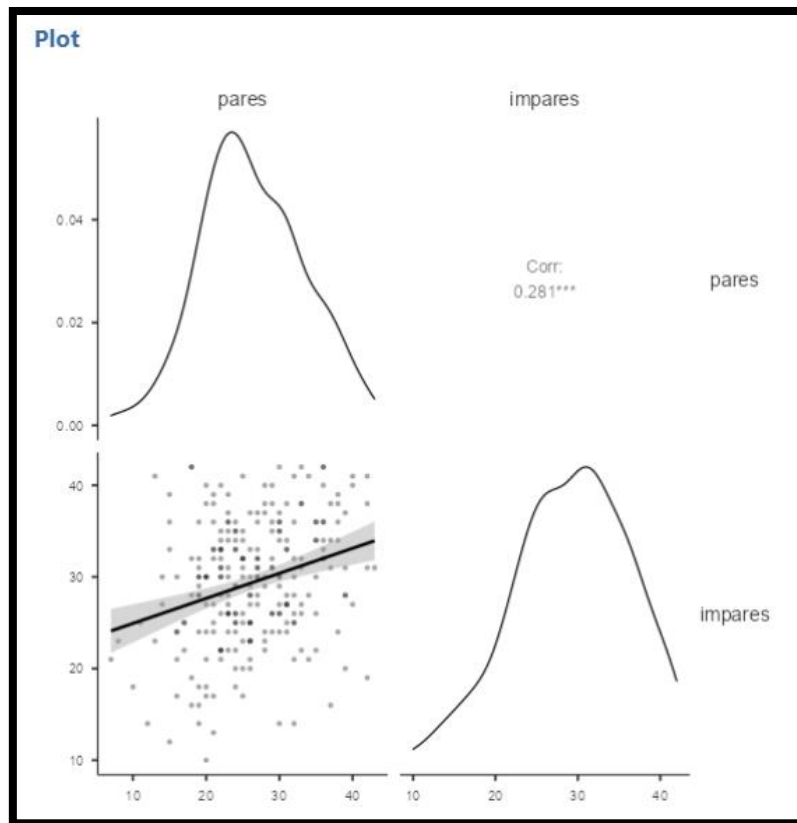
Plot:

- ☒ Correlation matrix
- ☒ Densities for variables
- ☒ Statistics

En este caso podemos observar que el resultado nos indica que la correlación entre ambas partes del test es positiva y significativa, pero de pequeña magnitud ($r_{xx} = 0,281$). Esta correlación nos indica la fiabilidad de cada una de las dos partes del test. Para conocer la fiabilidad de las puntuaciones totales del test, tenemos que aplicar la **fórmula de la profecía de Spearman-Brown**:

$$r_{pxx} = \frac{2 * r_{xx}}{1 + r_{xx}} = 0,438$$

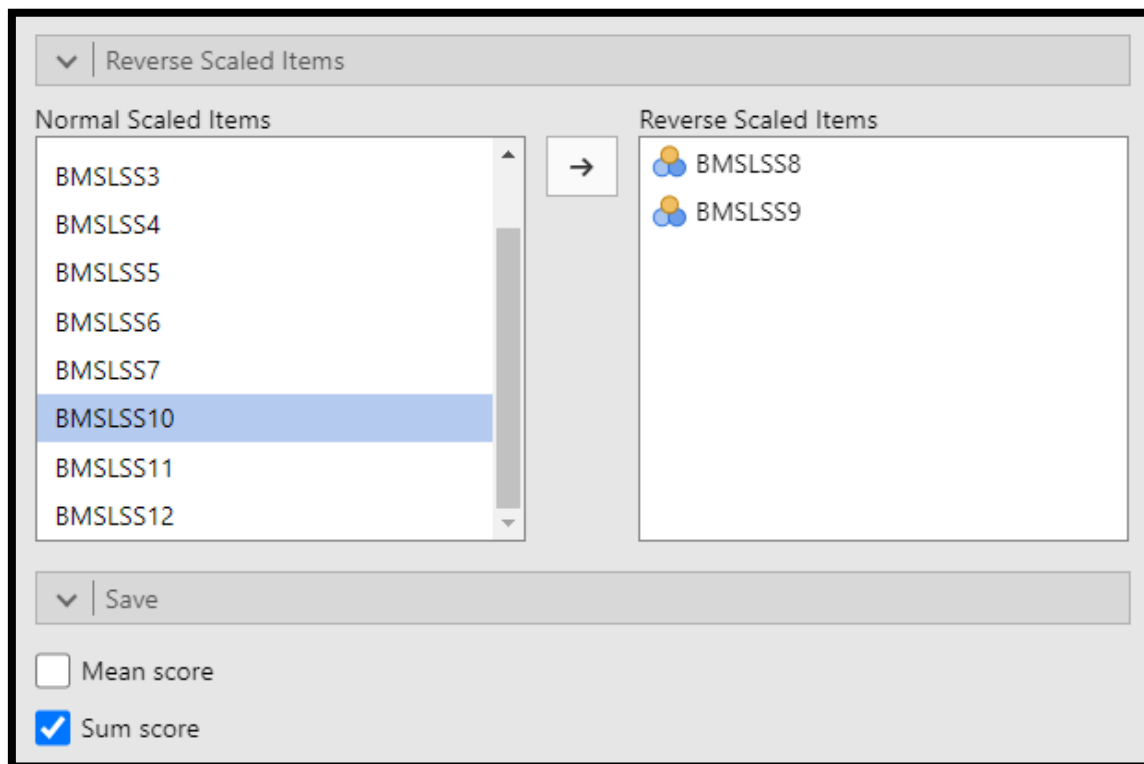
En este caso, concluiríamos que la fiabilidad de las puntuaciones totales de la escala medida por el método de dos mitades es **no adecuada**. Sin embargo, es importante recordar que esta aproximación a la fiabilidad es bastante limitada y tiende a no darnos un cálculo adecuado, ya que depende de la manera en la que nosotros hayamos decidido dividir el test.



1.4. Análisis de la fiabilidad por el método de test-retes

Para poder aproximar la **fiabilidad utilizando el método del test-retest** (también llamada *estabilidad o consistencia temporal*), es necesario haber medido a las mismas personas en los mismos ítems en, al menos, dos ocasiones. Para poder realizar su cálculo es necesario obtener las puntuaciones totales de la escala en ambas ocasiones. Para ello, podemos utilizar la función **Calcular variable** y sumar los ítems correspondientes en cada momento, o utilizar el menú de **Factor >> Reliability Analysis**.

Para obtener las puntuaciones sumadas de la escala BMSLSS en el primer momento, nos basta con indicar los ítems en el menú de *ítems*, recodificar los ítems necesarios, y abrir el menú **Sum Scores**. En este momento podemos pedir a Jamovi que nos calcule la puntuación sumada en la escala o la puntuación media. En el momento que hacemos click, se nos calculará una nueva variable con los ítems recodificados como le hemos indicado en el menú anterior.




Es conveniente revisar que Jamovi ha realizado la suma correctamente y darle un nombre que nos permita identificar rápidamente dicha variable.

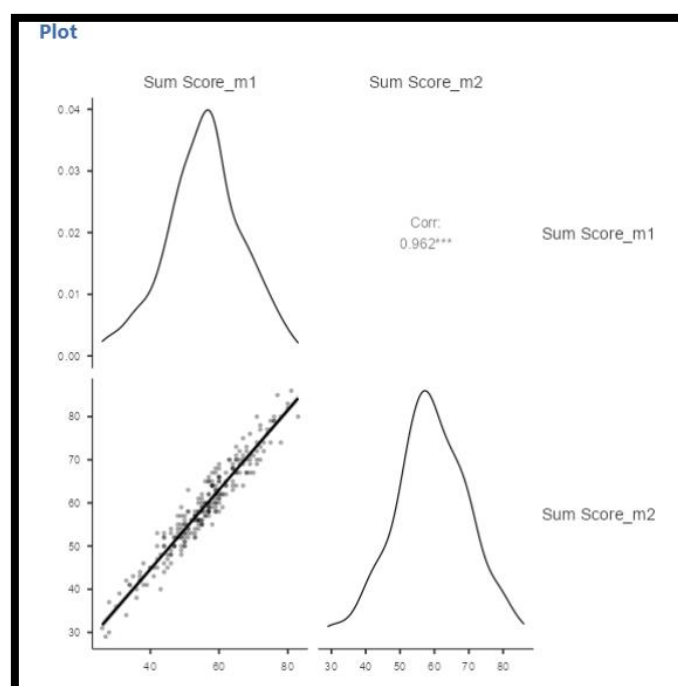
OUTPUT VARIABLE

Sum Score_m1

Sum score based on the variables 'BMSLSS1', 'BMSLSS2', 'BMSLSS3', 'BMSLSS4', 'BMSLSS5', 'BMSLSS6', 'BMSLSS7', 'BMSLSS8 (reversed)', 'BMSLSS9 (reversed)', 'BMSLSS10', 'BMSLSS11', and 'BMSLSS12'



Si abrimos un nuevo menú de **Analysis >>Factor >> Reliability Analysis**, podemos realizar el mismo proceso para las puntuaciones en el momento 2. Es importante que no sustituyamos las variables del análisis anterior, ya que la variable calculada se actualizaría y perderíamos las puntuaciones sumadas para el momento 1. Es necesario realizar dos veces el proceso. Una vez hayamos obtenido ambas puntuaciones, podemos calcular la fiabilidad test-retest como la correlación entre las puntuaciones en el test en ambos momentos. En este caso podemos observar que la consistencia temporal de las puntuaciones es muy alta ($r_{xx} = 0,962$).



En última instancia, podemos calcular la estabilidad temporal para las puntuaciones de cada uno de los ítems. Para ello, necesitamos calcular la correlación entre cada uno de los pares de ítems. Es recomendado hacer este cálculo ítem a ítem, ya que la tabla de correlaciones correspondiente es normalmente de un tamaño excesivo para su correcta interpretación. Por ejemplo, las puntuaciones del ítem 1 muestran una buena estabilidad temporal entre los dos momentos analizados.

Correlation Matrix			
		BMSLSS1	BMSLSS1_m2
BMSLSS1	Pearson's r	—	
	p-value	—	
BMSLSS1_m2	Pearson's r	0.877	—
	p-value	< .001	—

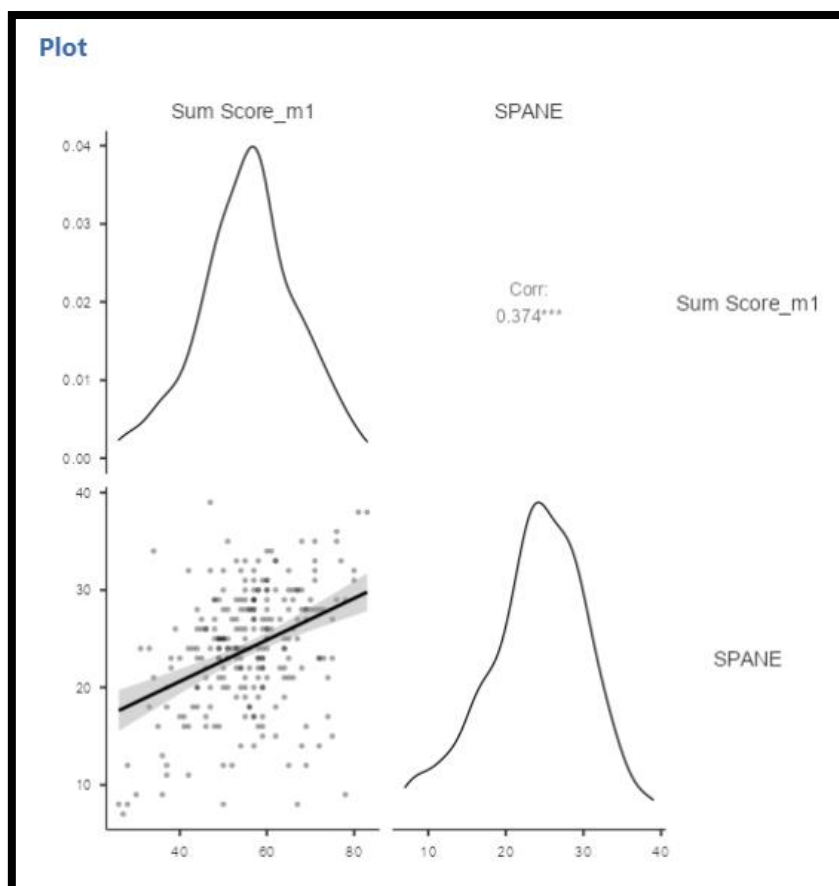
Por último, podríamos obtener el efecto de eliminar cada ítem sobre la fiabilidad test-retest. Para ello, tendríamos que calcular la puntuación sumada en el momento 1 y momento 2 **sin considerar** cada uno de los ítems, y ver como cambia dicho índice de fiabilidad. Este proceso tendría que repetirse para cada uno de los pares de ítems que componen la escala. Para facilitar el proceso, podemos volver al cálculo de la puntuación sumada en cada momento utilizando el menú de **Reliability Analysis**, eliminar el ítem correspondiente del menú de *ítem* y quitar la selección y seleccionar el que nos calcule la puntuación sumada para que se actualice dicho cálculo. Si calculamos la fiabilidad test-retest sin considerar el ítem 1, observamos que la misma disminuye una centésima. Esto nos indica que no deberíamos considerar eliminar dicho ítem.

Correlation Matrix			
		Sum Score_m1	Sum Score_m2
Sum Score_m1	Pearson's r	—	
	p-value	—	
Sum Score_m2	Pearson's r	0.961	—
	p-value	< .001	—

2: Evidencias de Validez

2.1. Análisis del índice o coeficiente de validez.

El índice o coeficiente de validez se define como la correlación entre las puntuaciones en el test y las puntuaciones en un criterio. En este caso, vamos a utilizar como criterio las puntuaciones en la subescala de emociones positivas del SPANE. Lo primero que necesitamos calcular para obtener el índice de validez es obtener la puntuación sumada para los ítems que conforman el criterio. Podemos utilizar el menú de **Analysis >> Factor >> Reliability Analysis** para obtener las puntuaciones totales en esta escala como hicimos previamente con las puntuaciones de nuestro test. Vamos a guardar dichas puntuaciones en una variable llamada **SPANE**. El índice de validez lo obtendríamos calculando la correlación entre las puntuaciones sumadas de nuestro test y la variable SPANE.



Es importante recordar que la correlación observada $r_{xy} = 0,374$ es la **correlación atenuada** entre las puntuaciones verdaderas de la subescala positiva de BMSLSS y SPANE. Podríamos calcular la correlación desatenuada utilizando los índices de fiabilidad de BMSLSS (ω de McDonald = 0,774) y SPANE (ω de McDonald = 0,660).

$$r_{vxy} = \frac{0,374}{\sqrt{0,774 * 0,660}} = 0,52$$

Por lo tanto, ha existido un **efecto de la atenuación** de $0,52 - 0,374 = 0,146$.

2.2. Análisis de validez incremental

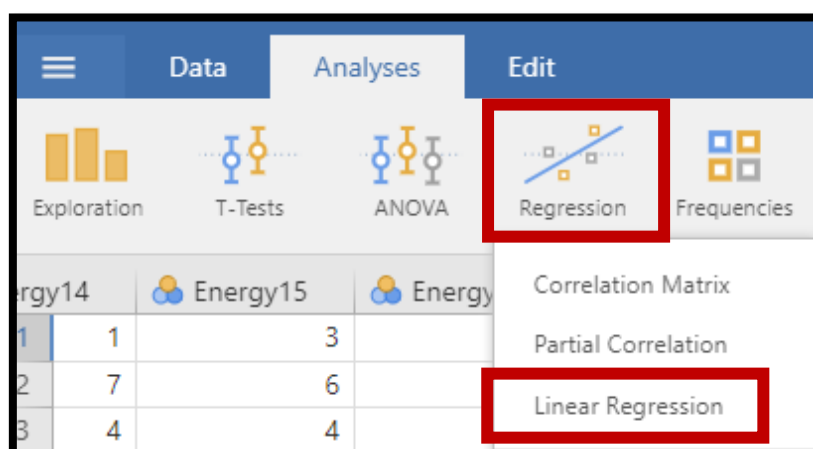
En este caso vamos a estar interesados en conocer la validez incremental de las puntuaciones de BMSLSS al predecir las puntuaciones de la escala positiva del SPANE una vez hemos considerado las puntuaciones de energía. El primer paso para realizar este análisis es calcular la puntuación en el test de **Energía/Activación**. Podemos utilizar el menú de **Analysis >> Factor >> Reliability Analysis** para obtener las puntuaciones totales en esta escala como hicimos previamente con las puntuaciones de nuestro test.

El primer paso para obtener evidencia de validez incremental es obtener los índices de validez y los **coeficientes de determinación** para cada una de las dos variables. En este caso, ya observamos un coeficiente de validez $r_{xy} = 0,374$ para BMSLSS. El coeficiente de determinación correspondiente sería $r^2_{xy} = 0,14$, lo que nos indica que las puntuaciones de bienestar predicen el 14% de la varianza de las puntuaciones de experiencias positivas.

Podemos obtener el coeficiente de validez para las puntuaciones de Energía/Activación de la misma manera (menú **Regression >> Correlation Matrix**). Observamos que la correlación entre ambas variables es igual a $r_{xy} = 0,608$, lo que nos indica un coeficiente de determinación $r^2_{xy} = 0,37$. Esto nos indica que las puntuaciones de Energía/Activación predicen el 37% de las puntuaciones de experiencias positivas.

Correlation Matrix			
		SPANE	Energy
SPANE	Pearson's r	—	
	p-value	—	
Energy	Pearson's r	0.608	—
	p-value	< .001	—

Para poder poner a prueba la validez incremental de las puntuaciones del test BMSLSS, necesitamos calcular dos modelos de regresión diferentes: a) un modelo donde predecimos las puntuaciones de emociones positivas a partir de las puntuaciones de Energía/Activación; y b) un modelo donde predecimos las puntuaciones de emociones positivas a partir de las puntuaciones de Energía/Activación **y BMSLSS**. Para calcular estos modelos, vamos a utilizar el menú de **Analysis >> Regression >> Linear Regression**.



En este menú vamos a incluir nuestra variable criterio dentro del cuadro de variable *Dependent*, y vamos a incluir la variable del primer modelo dentro del cuadro de variables *Covariates*. Si abrimos el menú **Model Builder**, veremos que la variable Energía/Activación se ha incluido en el primer bloque del análisis de regresión.

Linear Regression

Energy4

Energy5

Energy6

SPANE-P1

SPANE-P2

SPANE-P3

SPANE-P4

SPANE-P5

SPANE-P6

Autonomy

Negative Feelings

Sum Score_m1

→

→

→

Dependent Variable

SPANE

Covariates

Energy

Factors

Model Builder

Predictors

Energy

Blocks

Block 1

Energy

+ Add New Block

Los resultados del modelo de regresión confirman los resultados que vimos previamente en el análisis de correlación.

Model Fit Measures		
Model	R	R ²
1	0.608	0.370

Para añadir la variable BMSLSS en un nuevo modelo, vamos a hacer click en la opción **+ Add New Block** dentro del menú *Model Builder*. A continuación vamos a incluir la variable que recoge las puntuaciones sumadas del test BMSLSS en el cuadro de Covariates. Jamovi automáticamente incluirá la nueva variable dentro del segundo paso del análisis. Jamovi nos va a devolver los resultados de ajuste total (R^2) para el modelo 1 (incluyendo Energía/Activación) y modelo 2 (Energía/Activación + BMSLSS). Además, Jamovi nos devuelve un cuadro de comparación en ajuste, indicándonos que las puntuaciones de BMSLSS únicamente predicen un 2.6% adicional de la varianza de las puntuaciones en SPANE una vez que hemos considerado las puntuaciones de Energía/Activación. Además, nos proporciona una prueba de significación para saber si existe una diferencia significativa entre la varianza explicada entre ambos modelos.

Model Fit Measures						
Model	R	R^2				
1	0.608	0.370				
2	0.629	0.396				

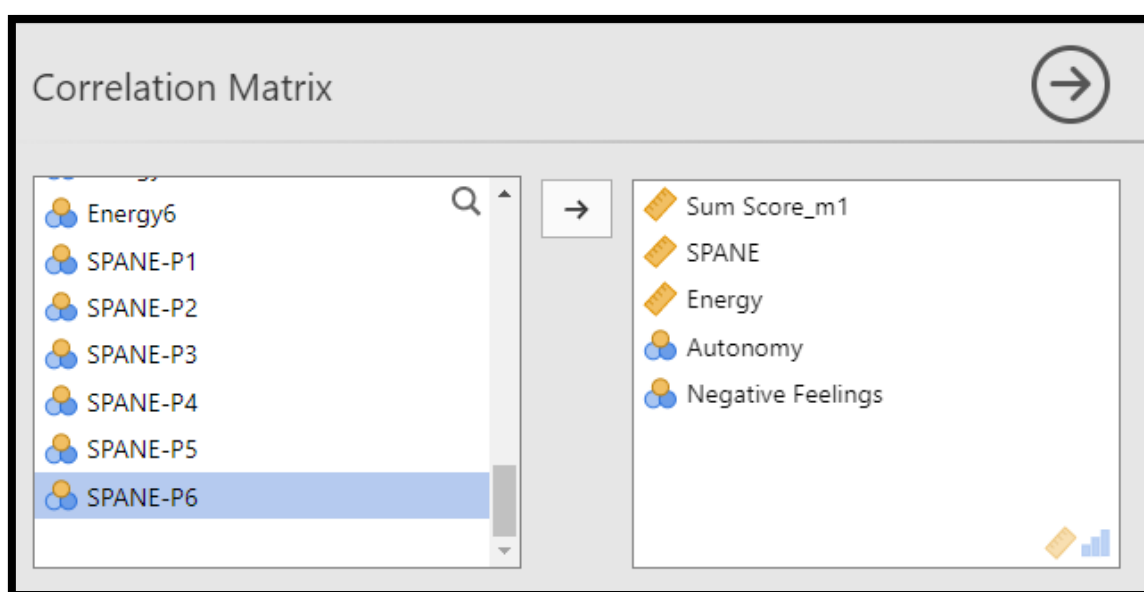
Model Comparisons						
Comparison						
Model	Model	ΔR^2	F	df1	df2	p
1	- 2	0.026	10.787	1	247	0.001

Si nos fijamos en los resultados previos, BMSLSS presentaba coeficiente de validez con SPANE de $r_{xy}=0,374$ (coeficiente de determinación correspondiente sería $r^2_{xy}=0,14$). En este caso, únicamente un 2.6% de ese 14% de varianza de SPANE predicha por BMSLSS es varianza incremental de dicho test (varianza en SPANE no predicha ya por el test de Energía/Activación¹).

¹ BMSLSS y Energía/Activación tienen una correlación igual a 0,368, lo que equivale a un coeficiente de determinación de 0,135. Si se calcula la correlación parcial entre SPANE y BMSLSS controlando por Energy se puede observar el valor exacto de la validez incremental de BMSLSS.

2.3. Análisis de validez congruente y discriminante

En este último análisis vamos a analizar la validez congruente y discriminante de las puntuaciones del test BMSLSS. Para ello, vamos a utilizar las puntuaciones ya sumadas de la subescala de emociones negativas del SPANE y Autonomía. Según la teoría previa, las puntuaciones en bienestar deberían estar positivamente correlacionadas con la subescala de emociones positivas del SPANE y la escala de Energy/Activation, negativamente con la escala de emociones negativas del SPANE y ser independientes de las puntuaciones de Autonomía.



Podemos obtener las correlaciones entre todas las variables utilizando el menú **Analysis >> Regression >> Correlation Matrix**. En este caso, podemos introducir las cinco variables en el cuadro de análisis para obtener todas las correlaciones que nos interesa analizar.

Correlation Matrix						
Correlation Matrix		Sum Score_m1	SPANE	Energy	Autonomy	Negative Feelings
Sum Score_m1	Pearson's r	—				
	p-value	—				
SPANE	Pearson's r	0.374	—			
	p-value	< .001	—			
Energy	Pearson's r	0.368	0.608	—		
	p-value	< .001	< .001	—		
Autonomy	Pearson's r	-0.118	0.095	-0.014	—	
	p-value	0.063	0.133	0.822	—	
Negative Feelings	Pearson's r	0.054	-0.051	-0.062	0.137	—
	p-value	0.392	0.419	0.331	0.031	—

Podemos observar que hemos observado evidencia de **validez convergente** en el caso de las correlaciones con la subescala de emociones positivas del SPANE y la escala de Energy/Activation. Por otra parte, **no** hemos observado que la correlación negativa esperada con las puntuaciones de la subescala de emociones negativas del SPANE, indicando una falta de **validez convergente** en este caso. Por último, hemos observado que las puntuaciones de nuestro test y Autonomía tienen una correlación de magnitud baja ($r = -0,118$), por lo que hemos obtenido **evidencia de validez discriminante**.

Si queremos ayudarnos de un gráfico para interpretar las correlaciones existentes entre las variables, podemos utilizar la opción gráfica incluida en **Analysis >> Factor >> Reliability Analysis**. Dentro del menú añadimos las variables al cuadro de *Ítems* e indicamos la opción del gráfico de **Additional Options (Correlation Heatmap)**.

Correlation Heatmap

