

Comparaciones por pares de medias en los diseños de medidas repetidas (ANOVA intra-sujetos)

Procedimiento de análisis

- Si el análisis del efecto estudiado es estadísticamente significativo y la variable tiene más de dos condiciones o mediciones ($A > 2$) será necesario **proseguir** con una prueba de contraste que permita conocer qué niveles de la variable de medidas repetidas difieren de forma estadísticamente significativa.
- Se aplicarán pruebas t de medidas repetidas para analizar la diferencia entre las medias de cada par de medias **+** una prueba de corrección o ajuste del α :
 - Holm–Bonferroni (Holm)
 - Šidák
 - Bonferroni
- No se utilizarán las Tukey HSD (incorrecta) y Scheffé (muy conservadora) para analizar las diferencias entre los pares de medias en diseños de medidas repetidas.

Comparaciones por pares de medias en diseños de medidas repetidas (ANOVA intra-sujetos)

Contexto: ANOVA intra-sujetos (diseño de medidas repetidas)

- **Qué prueba utilizar para los contrastes entre pares de medias:**

- Siempre: pruebas *t* para muestras relacionadas acompañadas de corrección de la tasa de error de Tipo I con alguna prueba estadística.

- **Por qué hay que ajustar α :**

- Porque al hacer varios contrastes aumenta la probabilidad de cometer más error de Tipo I del fijado a priori con el valor de alfa.
- Queremos controlar la **tasa de error Tipo I por experimento**.

- **Qué correcciones de α son apropiadas:**

- Bonferroni
- Šidák
- Holm–Bonferroni (Holm)

- **Esas 3 pruebas:**

- No requieren independencia estricta entre comparaciones.
- Mantienen bajo control la tasa de error de Tipo I en el conjunto de pruebas.

Comparaciones por pares de medias en los diseños de medidas repetidas (ANOVA intra-sujetos)

Comparación entre Bonferroni, Šidák y Holm–Bonferroni

Método	Ajuste del alfa por comparación $C = \text{nº de comparaciones}$	Ventajas	Inconvenientes
-Bonferroni	$\alpha_{\text{ajustada}} = \alpha / C$ (Nota ajuste lineal: si ajustamos el valor de p , entonces : $p_{\text{ajustado}} = p \sin \text{ajuste} \times C$)	-Muy robusto, muy fácil de entender y aplicar.	Más conservador (menos potente).
-Šidák	$\alpha_{\text{ajustada}} = 1 - (1 - \alpha)^{1/C}$ (Nota ajuste exponencial: si ajustamos el valor de p , entonces el exponente es C : $p_{\text{ajustado}} = 1 - (1 - p \sin \text{ajuste})^C$)	-Un poco menos conservador que Bonferroni.	Derivado bajo independencia; en la práctica: mejora pequeña.
-Holm–Bonferroni	Ordena los p y aplica Bonferroni secuencial	-Controla la tasa de error de Tipo I por experimento en sentido fuerte; -más potente que Bonferroni. -no requiere independencia estricta.	Ligeramente más complejo de explicar que Bonferroni simple. SPSS no lo calcula.

Comparaciones por pares de medias en los diseños de medidas repetidas (ANOVA intra-sujetos)

Comparación entre Bonferroni, Šidák y Holm–Bonferroni

Resumen:

- Las tres correcciones controlan adecuadamente la tasa de error Tipo I.
- **Holm–Bonferroni (Holm)** ofrece el **mejor equilibrio** entre control del error y potencia → opción preferente.
- **Šidák** y **Bonferroni** son alternativas válidas.
- Bonferroni es la **más conservadora** y la **más sencilla de aplicar**.

Comparaciones por pares de medias en los diseños de medidas repetidas (ANOVA intra-sujetos)

**Pruebas que no usaremos para pares en medidas repetidas:
Tukey y Scheffé**

Scheffé

- Es una prueba teóricamente correcta (controla la tasa de error de Tipo I por experimento para todos los contrastes lineales).
- Pero en diseños de medidas repetidas, cuando solo nos interesan analizar los contrastes entre los pares de medias:
 - Es muy conservadora.
 - Se pierde mucha potencia estadística frente a Holm, Šidák o Bonferroni.
- Uso recomendado:
 - Solo tendría sentido si queremos aplicar también contrastes complejos, no solo comparaciones por pares de medias.
- Indicación:
Scheffé no se utilizará para realizar comparaciones por pares en diseños de medidas repetidas.

Comparaciones por pares de medias en los diseños de medidas repetidas (ANOVA intra-sujetos)

**Pruebas que no usaremos para pares en medidas repetidas:
Tukey y Scheffé**

Tukey HSD

- Diseñada para:
 - ANOVA entre-sujetos,
 - con grupos independientes y varianzas homogéneas.
- En diseños de medidas repetidas:
 - Sus supuestos no se cumplen adecuadamente (especialmente cuando hay baja esfericidad).
 - Simulaciones (Maxwell, 1980) muestran que puede inflar la tasa de error Tipo I cuando se viola la esfericidad, incluso moderadamente.
- Por ello, se desaconseja la prueba de Tukey HSD como post hoc estándar en diseños de medidas repetidas (Field, 2013; Park et al., 2009; Maxwell et al., 2017).
- **Indicación:**
Nunca aplicaremos la prueba de Tukey HSD como si fuera un diseño entre-grupos cuando el diseño es de medidas repetidas.

INFORME DE INVESTIGACIÓN: SPSS

Base de datos. Diseño de medidas repetidas (ANOVA A = 3)

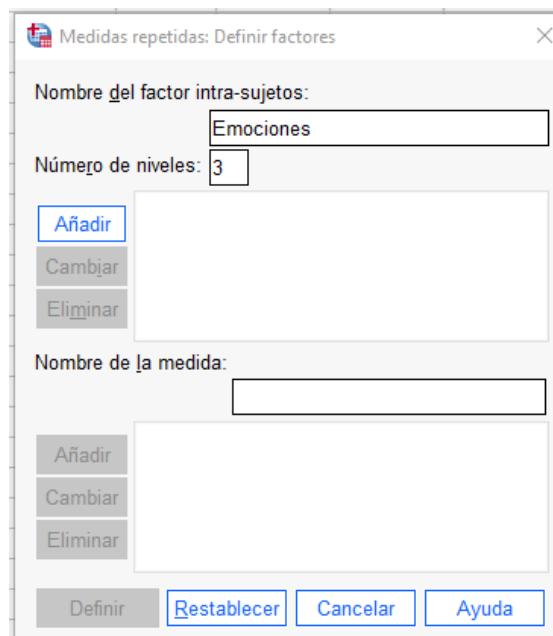
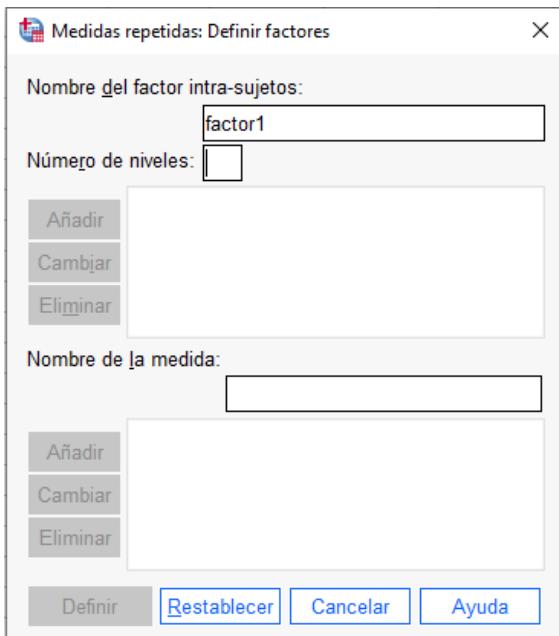
Variables: curiosidad, emociones negativas y emociones positivas

Comparaciones por pares de medias en SPSS (Bonferroni y Šidák y ningún ajuste)

En SPSS, las comparaciones por pares de medias de un factor intra sujetos se obtienen a partir del apartado de Medias marginales estimadas (**Estimated Marginal Means**):

1º Definir el factor intra-sujetos:

Menú: Analizar → Modelo lineal general → **Medidas repetidas** → Añadir → Definir



→ Y clicamos en **Añadir**
Y luego **Definir**

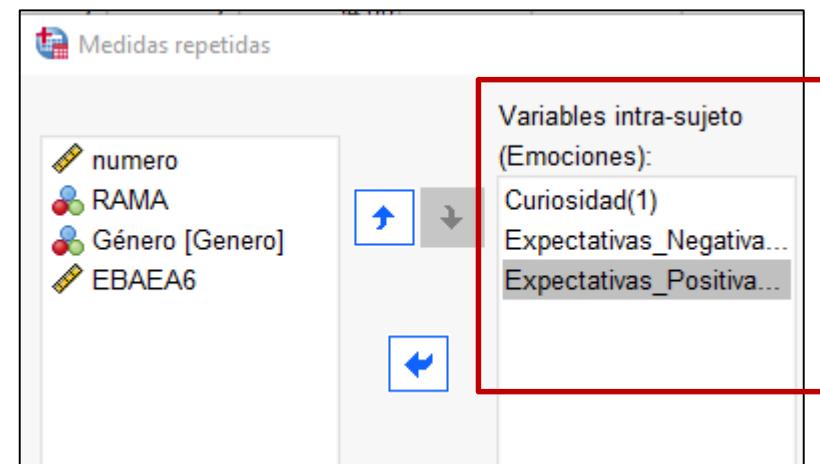
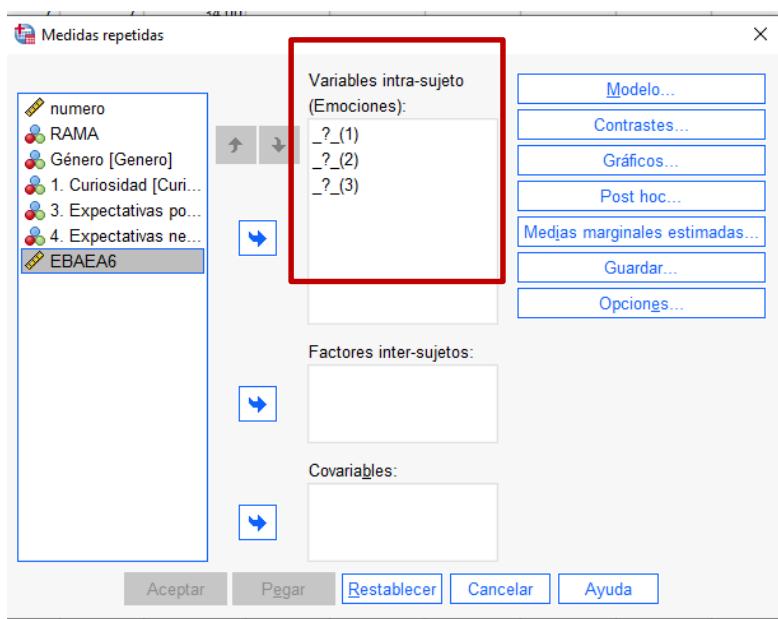
INFORME DE INVESTIGACIÓN: SPSS

Base de datos. Diseño de medidas repetidas (ANOVA A = 3)

Variables: curiosidad, emociones negativas y emociones positivas

2º Y añadimos las mediciones de la variable intra-sujeto completando los huecos de los interrogantes con cada variable:

curiosidad, expectativas negativas, expectativas positivas

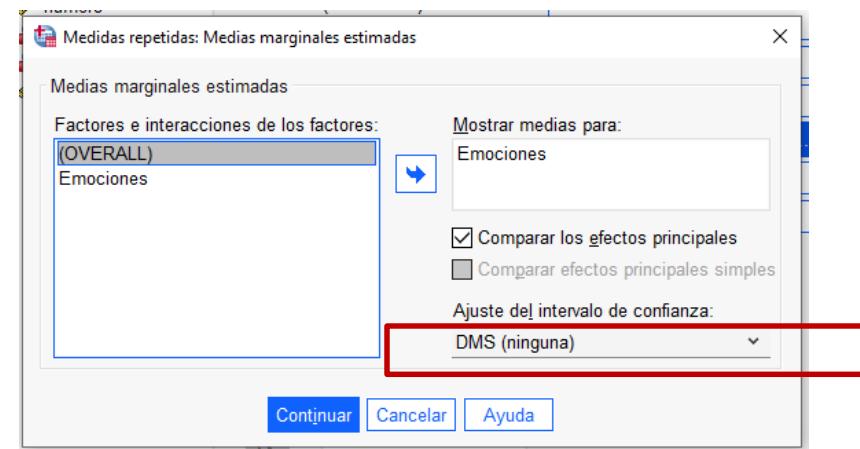
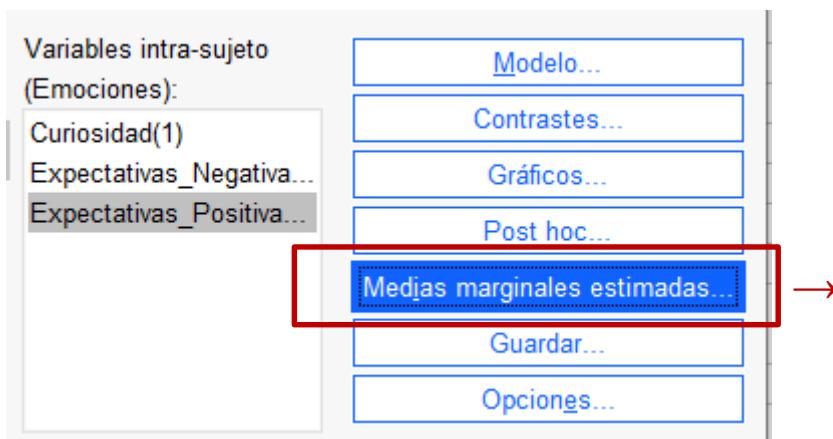


Base de datos. Diseño de medidas repetidas (ANOVA A = 3)

Variables: curiosidad, emociones negativas y emociones positivas

3º Seleccionar la prueba: DSM (ninguna) (sin ajustes):

Medias marginales estimadas → Comparar los efectos principales → Ajuste del intervalo de confianza → DSM (ninguna) (sin ajustes)



Base de datos. Diseño de medidas repetidas (ANOVA A = 3)

Variables: curiosidad, emociones negativas y emociones positivas

Medias marginales estimadas → DSM (ninguna) SIN AJUSTES → RESULTADOS

Comparaciones por parejas

Medida: MEASURE_1

(I) Emociones	(J) Emociones	Diferencia de medias (I-J)	Desv. Error	Sig. ^b	95% de intervalo de confianza para diferencia ^b	
					Límite inferior	Límite superior
1	2	-,119	,082	,149	-,281	,043
	3	2,124*	,158	<,001	1,812	2,435
2	1	,119	,082	,149	-,043	,281
	3	2,243*	,163	<,001	1,922	2,563
3	1	-2,124*	,158	<,001	-2,435	-1,812
	2	-2,243*	,163	<,001	-2,563	-1,922

Se basa en medias marginales estimadas

*. La diferencia de medias es significativa en el nivel ,05.

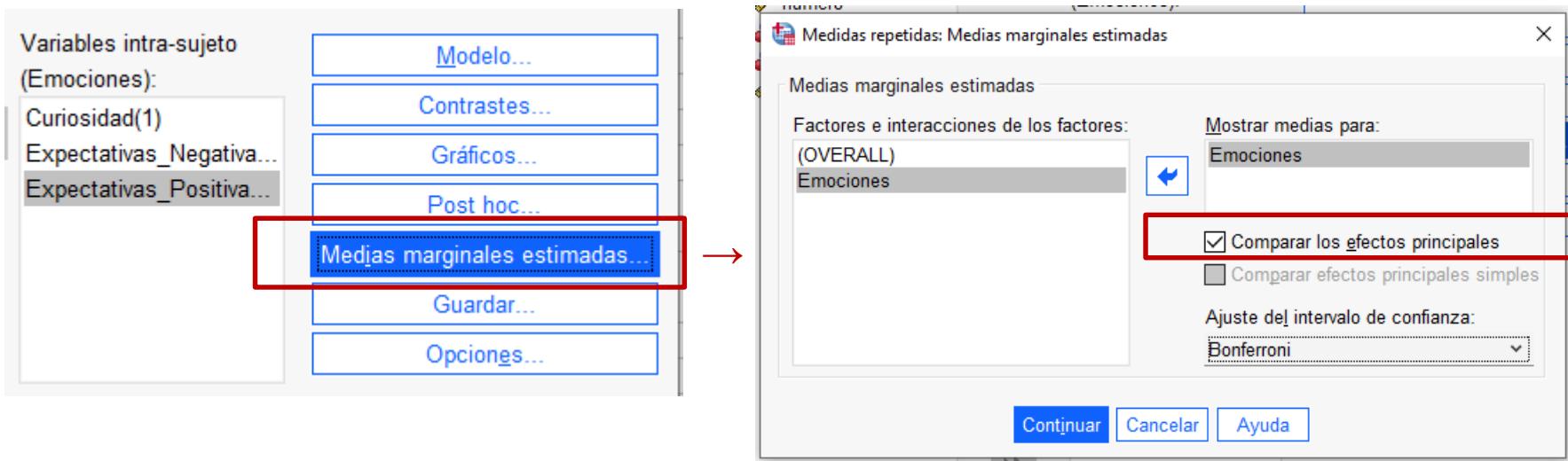
b. Ajuste para varias comparaciones: menor diferencia significativa (equivalente a sin ajustes).

Base de datos. Diseño de medidas repetidas (ANOVA A = 3)

Variables: curiosidad, emociones negativas y emociones positivas

3º Seleccionar la prueba: Bonferroni:

Medias marginales estimadas → Comparar los efectos principales → Ajuste del intervalo de confianza → BONFERRONI



Base de datos. Diseño de medidas repetidas (ANOVA A = 3)

Variables: curiosidad, emociones negativas y emociones positivas

Medias marginales estimadas → BONFERRONI → RESULTADOS

Comparaciones por parejas						
Medida: MEASURE_1	(I) Emociones	(J) Emociones	Diferencia de medias (I-J)	Desv. Error	Sig. ^b	95% de intervalo de confianza para diferencia ^b
	1	2	-,119	,082	,447	-,317 ,079
		3	2,124*	,158	<,001	1,743 2,505
	2	1	,119	,082	,447	-,079 ,317
		3	2,243*	,163	<,001	1,851 2,634
	3	1	-2,124*	,158	<,001	-2,505 -1,743
		2	-2,243*	,163	<,001	-2,634 -1,851

Se basa en medias marginales estimadas

*. La diferencia de medias es significativa en el nivel ,05.

b. Ajuste para varias comparaciones: Bonferroni.

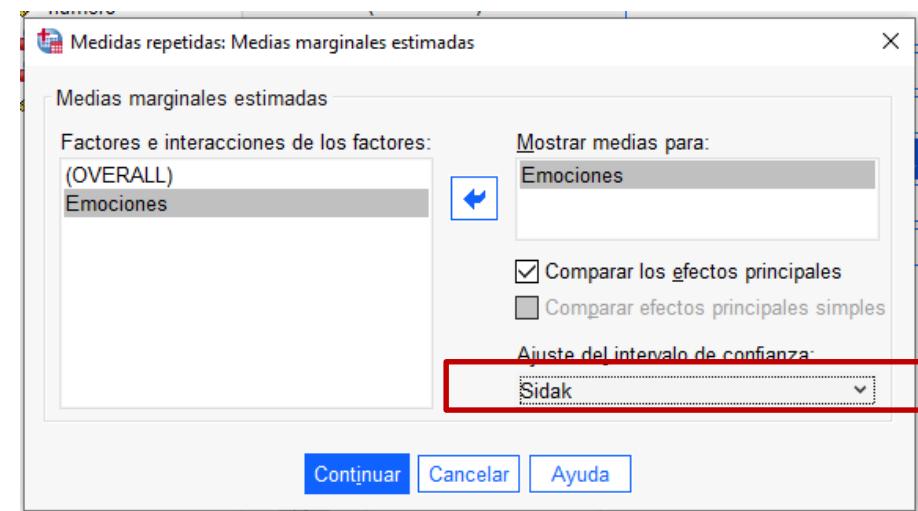
Nota: el **valor de *p* ajustado** de Bonferroni será aproximadamente ***p* sin ajuste x 3**. Por ejemplo: **a1 vs a2**: .149 x 3 = .45 y .447 aparece en el resultado de SPSS

Base de datos. Diseño de medidas repetidas (ANOVA A = 3)

Variables: curiosidad, emociones negativas y emociones positivas

3º Seleccionar la prueba: Sidak:

Medias marginales estimadas → Comparar los efectos principales → Ajuste del intervalo de confianza → SIDAK



Base de datos. Diseño de medidas repetidas (ANOVA A = 3)

Variables: curiosidad, emociones negativas y emociones positivas

Medias marginales estimadas

SIDAK

RESULTADOS

Comparaciones por parejas						
		Diferencia de medias (I-J)	Desv. Error	Sig. ^b	95% de intervalo de confianza para diferencia ^b	
(I) Emociones	(J) Emociones				Límite inferior	Límite superior
1	2	-,119	,082	,383	-,316	,078
	3	2,124*	,158	<,001	1,744	2,504
2	1	,119	,082	,383	-,078	,316
	3	2,243*	,163	<,001	1,852	2,633
3	1	-2,124*	,158	<,001	-2,504	-1,744
	2	-2,243*	,163	<,001	-2,633	-1,852

Se basa en medias marginales estimadas

*. La diferencia de medias es significativa en el nivel ,05.

b. Ajuste para varias comparaciones: Sidak.

Nota: Observar que Sidak ($p = .383$) es más potente que Bonferroni ($p = .447$). Ambas pruebas controlan la tasa de error de Tipo I.

No aplicar Tukey HSD ni Scheffé para comparaciones por pares en medidas repetidas porque están pensados para otros contextos y pueden ser inadecuados (Tukey) o demasiado conservadores (Scheffé).

Nota: el **valor de p ajustado por comparación** de Sidak será aproximadamente **p sin ajuste³**. En este caso se eleva cubo porque hay 3 comparaciones
Por ejemplo: **a1 vs a2**:

$$1 - (1 - .149)^3 = 1 - (.614) = .39$$
 y .383 aparece en el resultado de SPSS.
Un valor un pelín menor que el Bonferroni (que es más conservador), aunque con sólo 3 comparaciones la diferencia entre los dos ajustes es pequeña.

SPSS no ofrece: Holm–Bonferroni