

Diseño factorial: dos variables independientes, A x B

Hasta ahora, en el libro solamente se han considerado ejemplos donde el investigador ha sometido a contraste una sola variable independiente (un factor, diseños unifactoriales univariados), que podría tener dos condiciones, tres, cuatro o más. Mediante la aplicación de un diseño factorial, en cambio, se pretende analizar el efecto simultáneo de dos o más variables independientes (dos o más factores, diseños factoriales) sobre una variable dependiente medida (ver figura 47).



Figura 1. Diseño unifactorial y diseño factorial

Como ya se ha comentado anteriormente, las variables que explican los cambios que se producen en la variable dependiente son múltiples, y a veces es necesario analizar el efecto conjunto de dos o más factores para poder comprender convenientemente los cambios que se ocasionan en la variable dependiente.

Para definir un diseño factorial se numeran los factores que se analizan y el número de condiciones experimentales ('celdillas de interacción') que configuran cada uno de estos factores. Por ejemplo, el diseño factorial más simple es el 2×2 , es decir, el compuesto por dos *variables independientes* o factores con dos *niveles* cada una de ellas. Si se analizan dos factores, A con tres niveles o condiciones y B

con cuatro, se tratará de un diseño factorial 3 x 4. Cuando todos los factores tienen el mismo número de niveles suelen codificarse mediante una potencia: 2^3 es un diseño factorial de tres factores con dos niveles cada uno de ellos (representa lo mismo que $2 \times 2 \times 2$).

El número de condiciones experimentales resultantes de la aplicación de un determinado *diseño factorial* (las denominadas ‘celdillas de interacción’) será igual al producto de los niveles de los factores entre sí. En un diseño factorial 2×2 , existirán cuatro condiciones experimentales: a_1b_1 , a_1b_2 , a_2b_1 y a_2b_2 , resultado de combinar los dos niveles del factor A con los dos de B (ver figura 48).

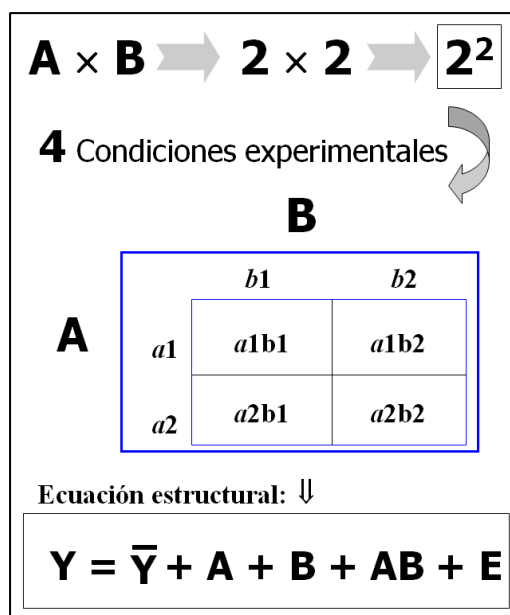


Figura 2. Tabla de datos de la interacción en un diseño 2×2

Ventajas del diseño factorial

La estrategia factorial tiene dos ventajas principales respecto a la unifactorial:

1) Aumenta la potencia estadística de la prueba de la razón F . Al analizar conjuntamente el efecto de dos o más variables independientes, si realmente estos factores están relacionados con la variable dependiente, el término residual será menor en el diseño factorial que si se analiza cada una de estas variables unifactorialmente. Por lo tanto, el diseño factorial es, en principio, más potente (la potencia estadística es mayor, $1 - \beta$) que varios diseños unifactoriales, puesto que cada uno de los factores (si realmente están relacionados con la variable dependiente) reduce el componente residual del modelo factorial o término de error

del modelo (denominador de la Razón F). Esa reducción del término de error se produce en el denominador del estadístico de la Razón F y por ello aumenta el valor de dicho estadístico, es decir, aumenta la varianza explicada o varianza atribuida a la fuente del efecto o tratamiento (numerador de la Razón F).

2) Plantea hipótesis teóricas de interacción entre las variables independientes. La segunda ventaja, ya no es únicamente cuantitativa en el sentido de aumentar la potencia estadística, sino que puede incluso variar la interpretación que se hace de la relación entre los factores y la variable dependiente. Si la relación que los factores tienen con la variable dependiente varía al manipularlos conjuntamente, quiere decir que el efecto de un factor depende de que el otro asuma ciertos valores. Es decir, lo importante no es valorar el efecto de cada factor de forma independiente sino explorar y valorar el efecto de interacción que se produce entre las condiciones de los factores sobre la puntuación obtenida en la variable dependiente. Cuando ocurre este fenómeno se produce un efecto en el modelo que se conoce con el término de *efecto de interacción*. Por el contrario, cuando en un diseño factorial se hace referencia al efecto de un factor individual (es decir, se analiza su efecto de forma independiente del resto de factores) se denomina *efecto principal* del factor en cuestión. Así, por ejemplo, en un diseño factorial A x B hay dos efectos principales (el de A y el de B) y un efecto de interacción (el de A x B). Cuando se aplica un diseño factorial, al investigador o investigadora realmente le interesa estudiar ese efecto de interacción (estará reflejado en la hipótesis del estudio), pero necesita estimar los dos efectos principales para poder estimar el efecto de interacción y por ello las tres fuentes de varianza son sometidas a contraste estadístico. Si el efecto de interacción no es estadísticamente significativo entonces se procederá con la interpretación de los efectos principales para comprobar si alguno de ellos o todos alcanzan un nivel de significación estadísticamente significativo ($p \leq \alpha$).

La complejidad de la conducta humana obliga a pensar que está determinada por múltiples causas conexas entre sí. Es por ello que el diseño factorial, que permite descubrir y analizar el sentido de las interacciones entre variables determinantes de la conducta, sea el diseño por excelencia y el de más amplio uso en el contexto de la Psicología experimental básica y aplicada.

Tipo de interacción entre los factores

El patrón interaccional es diverso en función de la relación que se detecte entre las variables o factores del modelo. Cuando existe interacción, las líneas se cruzan o convergen en algún punto, mientras que cuando no se produce el efecto de interacción, las líneas se mantienen paralelas, ya que la distancia entre las medias es constante. La representación gráfica de las puntuaciones medias de la interacción puede estar indicando un patrón de *interacción ordinal*, *interacción no ordinal* o *cruzada*, *interacción mixta* e *interacción no lineal* (véase Figura 11).

Una *interacción es ordinal* cuando el orden de superioridad de un factor sobre el otro se mantiene o es constante aunque el efecto cuantitativo puede variar. La interacción ordinal es *positiva* si se observa un crecimiento en el grupo mayor y una disminución en el menor y *negativa* si se produce un acercamiento entre los factores.

Un efecto de interacción se denomina *no ordinal* o *interacción cruzada* cuando el orden de superioridad entre los factores se cambia, no manteniéndose constante.

La *interacción mixta* es aquel efecto de interacción que siendo no ordinal presenta una tendencia clara hacia el cambio de orden, dando lugar a una interacción ordinal.

La *interacción no lineal* se caracteriza por la falta de vínculo lineal entre los factores.

La interacción ordinal y la no ordinal son las dos piezas claves del fenómeno de la interacción en los diseños factoriales.

La distinción entre los tipos de interacción es importante. Como [Lubin \(1961\)](#) señala, supongamos que un investigador está interesado en analizar el efecto de dos drogas y dos tipos de terapia en relación al éxito de la intervención efectuada (diseño entre-sujetos 2 x 2). Si la relación gráfica entre las dos variables no es paralela, pero nunca llegan a cruzarse las líneas se tratará de un efecto de *interacción ordinal*. En este caso, para cada intervención realizada el efecto de la droga tendrá el mismo orden de superioridad o inferioridad en relación al factor terapia. Así, si la droga A₁ tiene un efecto mayor sobre la terapia B₁ también lo tendrá sobre la terapia B₂. La droga A₂ tendrá un efecto menor en ambos tipos de

terapia. En este tipo de interacción ordinal las líneas representadas nunca se cruzarán.

Si se produce una interacción *no ordinal* o *cruzada* tiene que producirse un cruce entre dos o más factores. En este caso, los efectos de las drogas no mantienen el mismo orden para cada terapia. Puede ocurrir que la droga A₁ tenga un efecto mayor sobre la terapia B₁ pero menor sobre B₂ en comparación con la droga A₂. En definitiva, cuando existe un efecto estadísticamente significativo de interacción, la diferencia de efectos entre combinaciones factoriales no es constante.

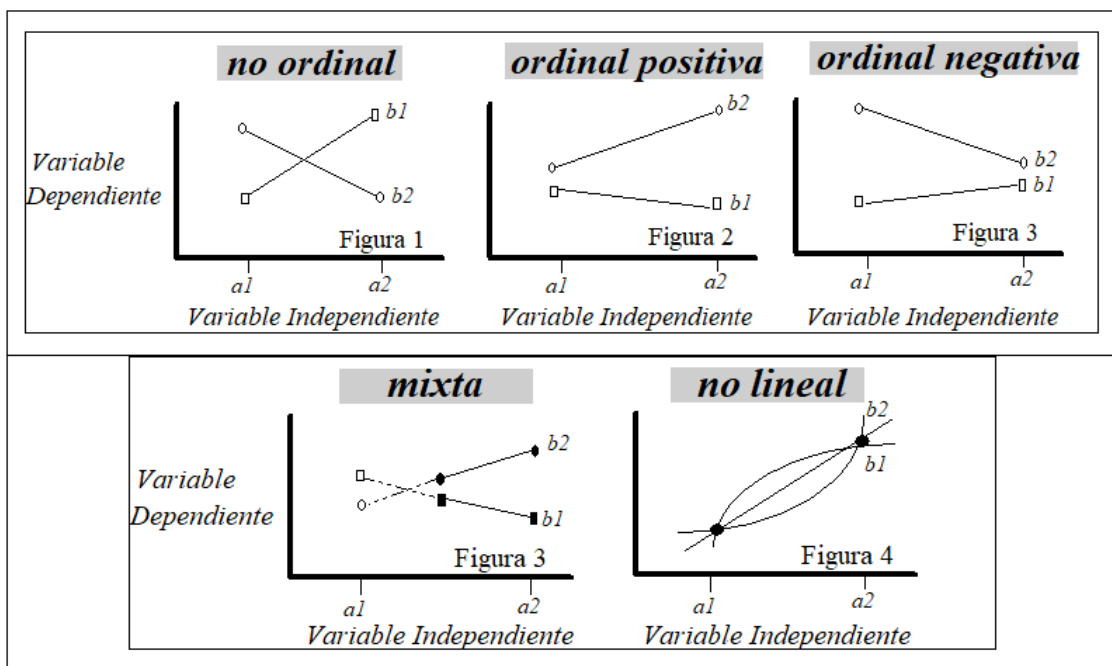


Figura 3. Representación del efecto de interacción

Modelo aditivo y no aditivo. Ecuación estructural

Para analizar el efecto conjunto o el efecto de interacción de dos factores A y B (se trata del diseño factorial más sencillo) sobre una única variable dependiente Y (diseño univariado) se define el *modelo completo* como un '*modelo de efectos no aditivos*':

$$Y = M + A + B + AB + E$$

Y = valores de la variable dependiente

M = media de la variable dependiente

A = efecto principal del primer factor, A

B = efecto principal del segundo factor, B

AB = efecto de interacción A × B

E = error de estimación del modelo

Podría ocurrir que en un diseño factorial se asumiese la no existencia del efecto de interacción entre los dos factores, ya que se estaría prediciendo la variable dependiente a partir de los efectos principales e independientes de los dos factores. En este caso se dice que se trata de un 'modelo de *efectos aditivos*' (sin el efecto de interacción entre los factores) cuya ecuación estructural es la siguiente:

$$Y = M + A + B + E$$

Y = valores de la variable dependiente

M = media de la variable dependiente

A = efecto principal del primer factor, A

B = efecto principal del segundo factor, B

E = error de estimación del modelo

El modelo de efectos aditivos se desarrollará posteriormente cuando se detalle el modelo de bloques con dos factores (un factor de tratamiento y un factor de bloqueo) cuya ecuación estructural no incluye el efecto de interacción ya que se plantea un diseño sin efecto de interacción entre las condiciones de los factores. Esta asunción de ausencia del efecto de interacción se conoce como 'aditividad del modelo' (efecto de interacción con $p > \alpha$; primer supuesto del diseño de bloques) y debe comprobarse antes de pasar a ejecutar el modelo factorial aditivo junto con comprobar también que el factor de bloqueo sí tiene un efecto estadísticamente significativo (efecto de la variable de bloqueo con $p < \alpha$; segundo supuesto del diseño de bloques). Si se detectará un efecto de interacción estadísticamente significativo entonces se habría producido un desajuste con la hipótesis teórica planteada ya que dicha hipótesis no hace referencia a un efecto de interacción entre los factores principales del modelo ya que se optó por este modelo aditivo.

En cambio, cuando el modelo incluye el término de interacción, que por definición es un componente multiplicativo, se dice que se está en presencia de un modelo de *efectos no aditivos*. En ambos casos, los dos modelos son aditivos en la medida que constan de componentes que sumados entre sí predicen la variable dependiente; pero sólo el modelo no aditivo contiene un término multiplicativo (efecto de interacción), entendido como un componente aditivo más del modelo de predicción.

Modelo no aditivo. Efecto de interacción

Cuando se trata de un modelo no aditivo (contiene el efecto de interacción), el punto crucial de la explicación de los resultados vendrá determinado por la significación estadística del componente de interacción entre los factores (A x B). Cuando se plantea este modelo no aditivo la hipótesis teórica plantea un efecto de interacción entre las variables independientes sobre la variable dependiente y, por ello, es la fuente de varianza que debe analizarse en primer lugar cuando ya se dispone de la tabla resumen de ANOVA.

Por lo tanto, cuando se plantea un diseño factorial no aditivo lo primero que hay que comprobar en la tabla de ANOVA es si el término de interacción es estadísticamente significativo y si lo es hay que centrar la interpretación en el análisis y reflexión de dicho efecto. Efectivamente, no tendría sentido centrar la explicación desde el efecto principal de la variable A (o de la variable B) cuando el efecto detectado en la variable dependiente depende de ciertas condiciones definidas por los valores de la otra variable (efecto de interacción).

La ecuación estructural de un modelo de diseño factorial con interacción (modelo no aditivo) se representa junto con la estimación de los efectos principales, interacción y el error en la Figura 50. Como se observa, la estimación del efecto de interacción consiste en restar de la media de la condición de interacción la media general y los efectos principales de los factores implicados en la interacción que se está analizando. Así, el efecto de la fuente de varianza de interacción entre los factores A y B se estima como:

$$AB = M_{ab} - M - A - B$$

Supongamos que se desea estimar el efecto de interacción de la celdilla a_1b_1 entonces habrá que calcular dicho efecto como:

$$A_1B_1 = M_{a_1b_1} - M - A_1 - B_1$$

Es decir, la estimación del efecto de interacción de la celdilla a_1b_1 es igual a la media de la interacción a_1b_1 menos la media general o constante, menos el efecto principal de la condición a_1 y menos el efecto principal de la condición b_1 . En este punto un error que se suele cometer al hacer la estimación del efecto de interacción es utilizar las medias de a_1 y b_1 y no los efectos de a_1 y b_1 . Por lo tanto, precaución cuando se calcula el efecto de interacción y recordar que se restan al final los efectos principales de las condiciones de cada factor que están implicadas en el efecto de interacción que se desea estimar.

En la figura 50 se observa que el término de error del modelo de diseño planteado siempre es la puntuación directa obtenida menos la puntuación predicha por el modelo formulado. Como se está presentando el diseño entre-grupos o entre-sujetos, se ha anotado que el error es $S / A \times B$, es decir, el error hace referencia a aquella parte de la variabilidad que se atribuye al error aleatorio, a las diferencias individuales de los participantes. En el diseño entre-grupos los sujetos están anidados (representado por la línea vertical, $S /$) en una cierta combinación de las condiciones del factor A y el factor B. De ahí que se represente el error como $S / A \times B$: sujetos anidados o ligados a una combinación concreta de los factores del diseño entre-grupos.

Teniendo en cuenta que la puntuación predicha siempre es la media general más los efectos que plantea el modelo de diseño en su ecuación estructural entonces en este modelo no aditivo $A \times B$, el error es igual a la puntuación obtenida en Y, menos la media general o constante, menos los efectos de A, de B y menos el efecto de la interacción $A \times B$. Conviene recordar que la puntuación predicha siempre es la media general o constante más los efectos que se plantean en la ecuación estructural (en este caso son A, B. y AB).

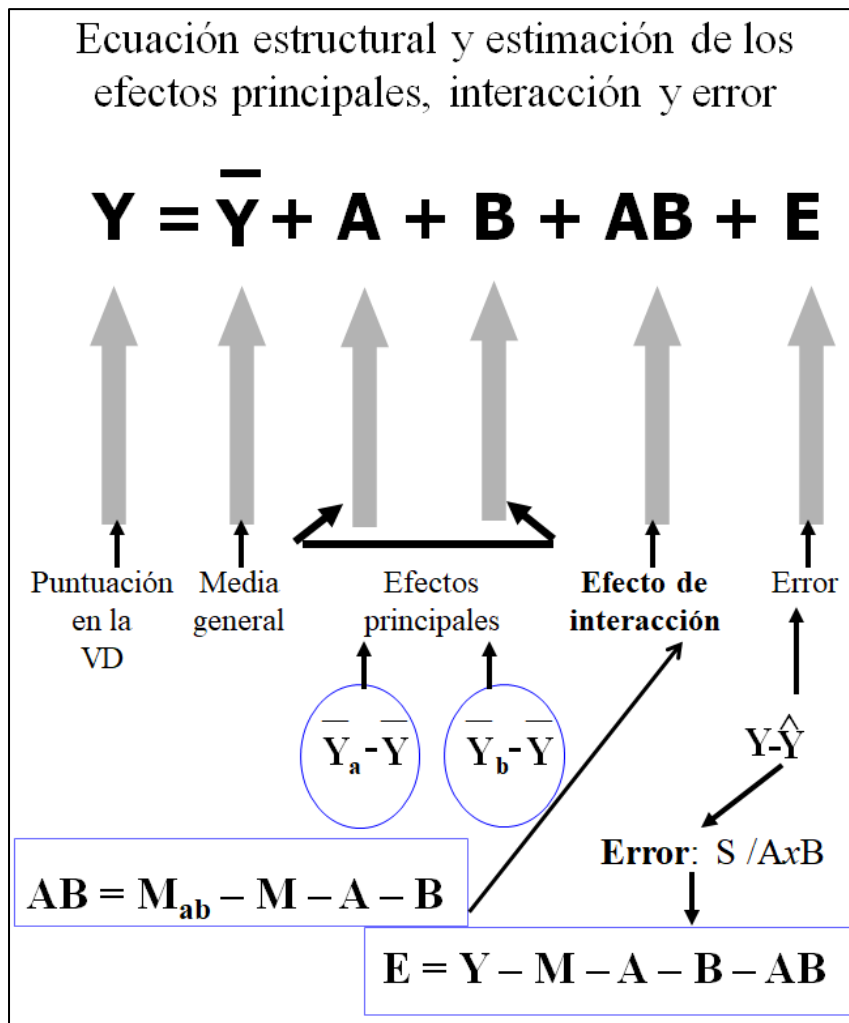


Figura 4. Ecuación estructural del diseño A x B y estimación de los efectos

A continuación se presenta un esquema con los conceptos básicos del modelo factorial no aditivo y aditivo tal y como ya se realizó anteriormente en el modelo de diseño unifactorial (ver Figura 51). Cada uno de los elementos que definen a los dos modelos (completo no aditivo y completo aditivo) se explica posteriormente con un ejercicio que plantea un supuesto de investigación.

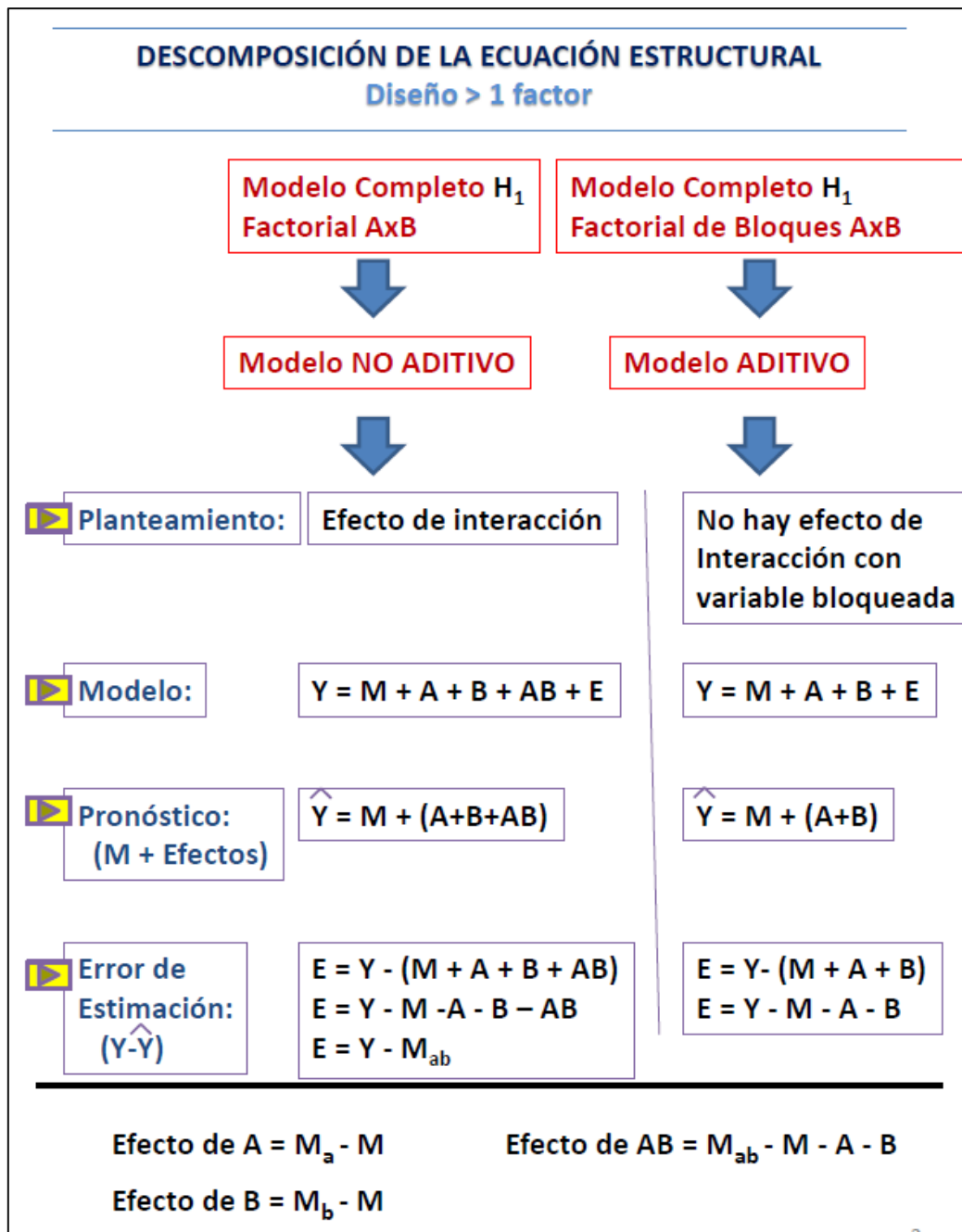


Figura 5. Descomposición de la ecuación estructural y modelos: aditivo y no aditivo

Desarrollo de un supuesto de investigación (A x B)

A continuación se plantea un supuesto de investigación con una hipótesis teórica de interacción entre dos variables independientes.

Supuesto de investigación. Un investigador desea comprobar si la memoria está relacionada con la motivación y el estrés percibido por los individuos. Para poder contrastar dicho supuesto, diseña una investigación en la que manipula

simultáneamente las variables de *Estrés* y *Motivación* (*bajo* y *alto*, en los dos factores). Antes de la prueba de recuerdo, la mitad de los participantes son sometidos a un cuestionario de preguntas matemáticas: en un caso las preguntas tienen una solución (estrés bajo) y en el otro grupo las preguntas no tienen una solución (y los sujetos son ‘ciegos’ a esta circunstancia, es decir, no saben que son preguntas sin una solución) y provoca que aumente el estrés de quienes se encuentran en esta condición ya que insisten e insisten en buscar una solución, pero no la hayan (estrés alto). Además, a la mitad de los sujetos se les ha notificado que recibirán una recompensa económica sustanciosa si resuelven adecuadamente la prueba de recuerdo de un listado de 20 palabras que posteriormente se les entregará (se trata de motivarles; motivación alta) y a la otra mitad no se les ha dicho nada (motivación baja). El diseño utiliza dos participantes en cada grupo experimental, asignándolos aleatoriamente a las condiciones experimentales. Los ocho participantes son varones seleccionados aleatoriamente de una facultad de economía de la Universidad de Navarra. Una vez que los participantes han sido sometidos a una de las condiciones de interacción del diseño (diseño entre-sujetos), los resultados en el número de errores cometidos al recordar el listado son los que se detallan en la tabla 11.

Tabla 1. Resultados del experimento y medias

		(A) <i>Estrés</i>		
		<i>a</i> ₁ <i>bajo</i>	<i>a</i> ₂ <i>alto</i>	<i>medias</i>
(B)	<i>b</i> ₁ <i>baja</i>	13	15	13
		11	13	
	Media <i>a</i> ₁ <i>b</i> ₁ : 12		Media <i>a</i> ₂ <i>b</i> ₁ : 14	
	<i>b</i> ₂ <i>alta</i>	1	11	7
3		13		
Media <i>a</i> ₁ <i>b</i> ₂ : 2		Media <i>a</i> ₂ <i>b</i> ₂ : 12		
<i>medias</i>		7	13	10

Una vez leído el supuesto de investigación es importante que el lector o lectora reflexione sobre el tipo de hipótesis de investigación planteada, el tipo de diseño de investigación que es apropiado para analizar dicha hipótesis, las variables dependientes e independientes que se utilizan y qué constructos representan y cómo se operacionalizan en el diseño, qué variables se han controlado y qué técnicas de control se han aplicado en el estudio (si las hay), la metodología que se

utiliza, el número de observaciones total y el número de observaciones por condición o grupo intraceldilla (¿el diseño es ortogonal o no?), qué valor de alfa se fija a priori como probabilidad del error de tipo I. Si en el texto del supuesto no se dice nada sobre el valor de alfa se asume que es .05 ya que si es otro valor, como por ejemplo $\alpha = .01$, sería necesario que el investigador o investigadora lo especificase en el texto del supuesto de la investigación de manera explícita. Del mismo modo, si en el texto del supuesto de investigación no se menciona nada sobre si el contraste de hipótesis es unidireccional o bidireccional entonces se asume que lo ha realizado de forma bilateral o a dos colas. Si fuese de una cola o unidireccional (la hipótesis marca claramente la dirección de la diferencia entre las medias) entonces es necesario que el investigador o investigadora lo resalte de forma explícita cuando describe el diseño de su estudio. A veces se plantean hipótesis bidireccionales y el investigador o investigadora marca la dirección de la diferencia según el conocimiento previo que ha revisado, pero desea aplicar hipótesis bidireccionales ya que no renuncia a observar también qué ocurre si la dirección de la diferencia es la contraria. En este caso, la elección del análisis es llevar a cabo el contraste de hipótesis estadísticas observando las dos colas de la distribución de la hipótesis nula ya que la diferencia podría ir en una dirección (grupo primero mayor que el segundo) o al revés (grupo segundo mayor que el primero). En general, en Psicología, en la mayoría de las ocasiones, se lleva a cabo el análisis con pruebas bilaterales y los investigadores o investigadoras no suelen mencionarlo en sus artículos ya que se asume que si no se dice nada esa es la forma como han actuado cuando han tomado la decisión estadística.

El primer paso para poder realizar el análisis de la varianza es conocer las puntuaciones medias de cada una de las condiciones de los efectos principales y las medias de interacción de las celdillas de interacción que se producen al cruzar las condiciones de los factores del diseño, situando además la media general o constante (tabla 12). En la tabla anterior ya se habían incluido, pero se detallan ahora para recordar la importancia de calcular las medias como primer paso para estimar posteriormente cada uno de los efectos que tiene el modelo. En la tabla de medias se puede observar la dirección de las medias. Así se comprueba que los sujetos que han recibido la condición de estrés bajo y motivación alta son los que menos palabras recuerdan. De momento las puntuaciones siguen el planteamiento

de la hipótesis teórica, pero falta comprobar si las medias difieren de forma estadísticamente significativa, y si difieren todos los pares de medias o solamente algunos de ellos. Para ello es necesario ejecutar un análisis de varianza con el proceso de decisión estadística.

Tabla 2. Tabla de medias

		(A) Estrés		medias
		a_1 bajo	a_2 alto	
Motivación	b_1 baja	12	14	13
	b_2 alta	2	12	7
medias		7	13	10

Por lo tanto, la ecuación estructural que habrá que descomponer para el supuesto de investigación planteado es la siguiente:

$$Y = M + A + B + AB + E$$

Efectos principales

Se pueden estimar los efectos principales del *bajo* y *alto* nivel de estrés (factor A) sobre el número de errores cometidos al recordar el listado a través de:

$$A = M_a - M = \begin{pmatrix} 7 \\ 7 \\ 7 \\ 7 \\ 13 \\ 13 \\ 13 \\ 13 \end{pmatrix} - \begin{pmatrix} 10 \\ 10 \\ 10 \\ 10 \\ 10 \\ 10 \\ 10 \\ 10 \end{pmatrix} = \begin{pmatrix} -3 \\ -3 \\ -3 \\ -3 \\ 3 \\ 3 \\ 3 \\ 3 \end{pmatrix} \rightarrow \begin{pmatrix} a_1 \\ a_1 \\ a_1 \\ a_1 \\ a_2 \\ a_2 \\ a_2 \\ a_2 \end{pmatrix}$$

El efecto de $\hat{\alpha}_1$ es de -3 puntos y el de $\hat{\alpha}_2$ de 3 puntos. Se comprueba que la suma de los efectos es cero. Por tanto, en la muestra se aprecia que el grupo sometido a un *estrés alto* aumenta su promedio de errores al recordar el listado en 3 puntos respecto de la media general, mientras que los que están en la condición de *bajo estrés* reducen tres puntos su puntuación promedio en el conteo de los errores cometidos al recordar el listado.

Los grados de libertad correspondientes al término del efecto principal A serán igual a 1 ($gl_A = a - 1 = 2 - 1 = 1$).

Se puede observar igualmente cómo afecta la motivación (factor B) a la cantidad de errores cometidos al recordar las palabras del listado:

$$\mathbf{B} = \mathbf{M}_b - \mathbf{M} = \begin{pmatrix} 13 \\ 13 \\ 7 \\ 7 \\ 13 \\ 13 \\ 7 \\ 7 \end{pmatrix} - \begin{pmatrix} 10 \\ 10 \\ 10 \\ 10 \\ 10 \\ 10 \\ 10 \\ 10 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ -3 \\ -3 \\ 3 \\ 3 \\ -3 \\ -3 \end{pmatrix} \rightarrow \begin{pmatrix} b_1 \\ b_1 \\ b_2 \\ b_2 \\ b_1 \\ b_1 \\ b_2 \\ b_2 \end{pmatrix}$$

Los cuatro casos en que la motivación es *baja* (b_1), aumentan los errores en el recuerdo de las palabras del listado en promedio 3 puntos ($\hat{\beta}_1 = 3$) respecto a la media general de errores de recuerdo cometidos. Si la motivación es *alta*, entonces disminuyen los errores en el recuerdo del listado en otros 3 puntos ($\hat{\beta}_2 = -3$) sobre el promedio general. Se observa de nuevo que la suma de los efectos del factor B es cero.

Los grados de libertad correspondientes al término del efecto principal B serán igual a 1 ($g_B = b - 1 = 2 - 1 = 1$).

Efectos de interacción en el modelo no aditivo

Para estimar el efecto de interacción AB se resta de la media de cada celdilla de interacción la media general o constante y los efectos principales de los factores que se han definido en la ecuación estructural; en este diseño son A y B.

Por lo tanto:

$$AB = M_{ab} - (\mathbf{M} + \text{EFECTOS PLANTEADOS EN EL MODELO O ECUACIÓN ESTRUCTURAL})$$

Es decir,

$$AB = M_{ab} - (\mathbf{M} + \mathbf{A} + \mathbf{B})$$

De lo que se deduce directamente que:

$$AB = M_{ab} - \mathbf{M} - \mathbf{A} - \mathbf{B}$$

Para obtener los efectos de interacción a partir de los datos del supuesto que se está desarrollando:

$$AB = M_{ab} - \mathbf{M} - \mathbf{A} - \mathbf{B} =$$

$$= \begin{pmatrix} 12 \\ 12 \\ 2 \\ 2 \\ 14 \\ 14 \\ 12 \\ 12 \end{pmatrix} - \begin{pmatrix} 10 \\ 10 \\ 10 \\ 10 \\ 10 \\ 10 \\ 10 \\ 10 \end{pmatrix} - \begin{pmatrix} -3 \\ -3 \\ -3 \\ -3 \\ 3 \\ 3 \\ 3 \\ 3 \end{pmatrix} - \begin{pmatrix} 3 \\ 3 \\ -3 \\ -3 \\ 3 \\ 3 \\ -3 \\ -3 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ -2 \\ -2 \\ -2 \\ -2 \\ 2 \\ 2 \end{pmatrix} \rightarrow \begin{pmatrix} ab_{11} \\ ab_{11} \\ ab_{12} \\ ab_{12} \\ ab_{21} \\ ab_{21} \\ ab_{22} \\ ab_{22} \end{pmatrix}$$

Los grados de libertad correspondientes al término de interacción se calculan multiplicando los grados de libertad de los efectos principales implicados; en este caso los grados de libertad serán igual a 1 ($g_{AB} = g_A g_B = 1 \times 1 = 1$).

Siguiendo con el análisis del efecto de interacción, ahora habrá que comprobar si estas diferencias entre las celdillas de interacción son o no estadísticamente significativas. Para ello se procede con la estimación de la fuente de varianza del error y, posteriormente, se completa el análisis de la varianza (ANOVA) con el estadístico de la razón F . En este dicho con dos efectos principales y un efecto de interacción será necesario calcular 3 razones F con sus correspondientes valores de p de probabilidad del efecto detectado (o un efecto mayor), asumiendo que la hipótesis nula es cierta; es decir, una prueba F para cada fuente de varianza y se lleva a cabo la decisión dicotómica de mantener o rechazar la hipótesis nula en los tres casos.

Error de estimación

El error de estimación que se comete está definido por la diferencia entre la puntuación obtenida en el estudio y la puntuación pronosticada por el modelo de diseño planteado en el estudio.

Respecto a la puntuación pronosticada, en un diseño A x B la estimación del valor de la variable dependiente se hace a partir de la suma de la media general \mathbf{M} y los efectos de \mathbf{A} , \mathbf{B} y \mathbf{AB} que se han planteado en la ecuación estructural. Por lo tanto, $\hat{\mathbf{Y}} = \mathbf{M} + \mathbf{A} + \mathbf{B} + \mathbf{AB}$.

Se puede analizar cuál será el pronóstico de cada una de las cuatro situaciones experimentales que tiene el diseño A x B del estudio planteado en el supuesto de investigación:

Estrés *bajo* / Motivación *baja* (a_1b_1):

$$\hat{Y} = \mathbf{M} + \mathbf{A} + \mathbf{B} + \mathbf{AB} = \bar{Y} + \hat{\alpha}_1 + \hat{\beta}_1 + \hat{\alpha\beta}_{11} =$$

$$= 10 + -3 + 3 + 2 = 12$$

Estrés *bajo* / Motivación *alta* (a_1b_2):

$$\hat{Y} = \mathbf{M} + \mathbf{A} + \mathbf{B} + \mathbf{AB} = \bar{Y} + \hat{\alpha}_1 + \hat{\beta}_2 + \hat{\alpha\beta}_{12} =$$

$$= 10 + -3 + -3 + -2 = 2$$

Estrés *alto* / Motivación *baja* (a_2b_1):

$$\hat{Y} = \mathbf{M} + \mathbf{A} + \mathbf{B} + \mathbf{AB} = \bar{Y} + \hat{\alpha}_2 + \hat{\beta}_1 + \hat{\alpha\beta}_{21} =$$

$$= 10 + 3 + 3 + -2 = 14$$

Estrés *alto* / Motivación *alta* (a_2b_2):

$$\hat{Y} = \mathbf{M} + \mathbf{A} + \mathbf{B} + \mathbf{AB} = \bar{Y} + \hat{\alpha}_2 + \hat{\beta}_2 + \hat{\alpha\beta}_{22} =$$

$$= 10 + 3 + -3 + 2 = 12$$

Se observa que cuando el modelo de diseño es un diseño entre-sujetos factorial no aditivo A x B, la puntuación pronosticada para cada uno de los cuatro efectos de interacción es la media de interacción de cada grupo. Sin embargo, conviene tener presente que eso sucede solamente en este diseño, pues si el diseño se plantea como aditivo entonces ya no se produce esa coincidencia tal y como se detallará después.

Siguiendo con los datos del supuesto de investigación, se observa que los participantes del grupo de estrés bajo y motivación alta (a_1b_2) son los que tienen menos errores al recordar el listado de las palabras. Y cuando el estrés es alto y la motivación es baja se produce el mayor número de errores al recordar el listado de las palabras.

Por lo tanto, el error de estimación que se comete para cada una de las observaciones que se han recogido en el estudio es el siguiente:

$$\mathbf{E} = \mathbf{Y} - \mathbf{M} - \mathbf{A} - \mathbf{B} - \mathbf{AB} =$$

$$= \begin{pmatrix} 13 \\ 11 \\ 1 \\ 3 \\ 15 \\ 13 \\ 11 \\ 13 \end{pmatrix} - \begin{pmatrix} 10 \\ 10 \\ 10 \\ 10 \\ 10 \\ 10 \\ 10 \\ 10 \end{pmatrix} - \begin{pmatrix} -3 \\ -3 \\ -3 \\ -3 \\ 3 \\ 3 \\ 3 \\ 3 \end{pmatrix} - \begin{pmatrix} 3 \\ 3 \\ -3 \\ -3 \\ -3 \\ -3 \\ 3 \\ 3 \end{pmatrix} - \begin{pmatrix} 2 \\ 2 \\ -2 \\ -2 \\ -2 \\ -2 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \\ 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} S1ab_{11} \\ S2ab_{11} \\ S3ab_{12} \\ S4ab_{12} \\ S5ab_{21} \\ S6ab_{21} \\ S7ab_{22} \\ S8ab_{22} \end{pmatrix}$$

Se observa de nuevo que la suma de todos los errores de estimación es igual a cero. También se puede comprobar que dentro de una celdilla concreta la suma de los errores estimados también debe dar cero. Es decir, si hay dos participantes que están en la misma condición de tratamiento (por ejemplo, a_1b_1), por qué sus puntuaciones directas difieren (13 y 11) si tienen la misma media general, el mismo efecto de A (a_1) y el mismo efecto de B (b_1), pues por las diferencias individuales y/o error aleatorio y de ahí que el valor del término de error es propio y vinculado a cada puntuación directa. Y el sumatorio de los errores dentro de cada celdilla (errores intra-celdilla) debe ser igual a cero si el diseño es ortogonal. Es decir, si se trata de un diseño ortogonal ($n_{a_1b_1} = n_{a_1b_2} = n_{a_2b_1} = n_{a_2b_2}$) entonces los grados de libertad del error se pueden calcular como $(n-1)ab$, siendo n el número de observaciones dentro de cada celdilla de interacción. Si el diseño no es ortogonal entonces los grados de libertad del error se pueden calcular como $N-ab$, siendo N el número total de observaciones. Las dos formulas de los grados de libertad ofrecen el mismo resultado, por supuesto.

En resumen, el desarrollo de la ecuación estructural (modelo no aditivo) para cada una de las observaciones del estudio es la siguiente:

$$\mathbf{Y} = \mathbf{M} + \mathbf{A} + \mathbf{B} + \mathbf{AB} + \mathbf{E}$$

$$\begin{pmatrix} 13 \\ 11 \\ 1 \\ 3 \\ 15 \\ 13 \\ 11 \\ 13 \end{pmatrix} = \begin{pmatrix} 10 \\ 10 \\ 10 \\ 10 \\ 10 \\ 10 \\ 10 \\ 10 \end{pmatrix} + \begin{pmatrix} -3 \\ -3 \\ -3 \\ -3 \\ 3 \\ 3 \\ 3 \\ 3 \end{pmatrix} + \begin{pmatrix} 3 \\ 3 \\ -3 \\ -3 \\ -3 \\ -3 \\ 3 \\ 3 \end{pmatrix} + \begin{pmatrix} 2 \\ 2 \\ -2 \\ -2 \\ -2 \\ -2 \\ 2 \\ 2 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \\ 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$$

Sumas de cuadrados

Una vez ya se dispone de la información vinculada a la descomposición de la ecuación estructural en cada uno de sus componentes, ya se procede con el cálculo de las sumas de cuadrados (SC) de los efectos de cada fuente de varianza planteada en el modelo de diseño de investigación: A, B, AB y E.

Suma de cuadrados del efecto de A, nivel de estrés:

$$\mathbf{SC}_A = \mathbf{A}' \mathbf{A} = \begin{pmatrix} -3 & -3 & -3 & -3 & 3 & 3 & 3 & 3 \end{pmatrix} \begin{pmatrix} -3 \\ -3 \\ -3 \\ -3 \\ 3 \\ 3 \\ 3 \\ 3 \end{pmatrix} = 72$$

Suma de cuadrados del efecto de B, nivel de motivación:

$$\mathbf{SC}_B = \mathbf{B}' \mathbf{B} = \begin{pmatrix} 3 & 3 & -3 & -3 & 3 & 3 & -3 & -3 \end{pmatrix} \begin{pmatrix} 3 \\ 3 \\ -3 \\ -3 \\ 3 \\ 3 \\ -3 \\ -3 \end{pmatrix} = 72$$

Suma de cuadrados del efecto de interacción AB, nivel de estrés por nivel de motivación:

$$\mathbf{SC}_{AB} = \mathbf{AB}' \mathbf{AB} = \begin{pmatrix} 2 & 2 & -2 & -2 & -2 & -2 & 2 & 2 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \\ -2 \\ -2 \\ -2 \\ -2 \\ 2 \\ 2 \end{pmatrix} = 32$$

Suma de cuadrados de la fuente de varianza del error, E:

$$SC_E = E' E = \begin{pmatrix} 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \\ 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} = 8$$

A continuación, se aplica la prueba de la hipótesis para determinar si la variabilidad correspondiente al término de interacción permite desechar el modelo aditivo y mantener la validez de un modelo con término de interacción o modelo no aditivo (tabla 13). Se trata de la prueba de aditividad del modelo. Previamente se dividen las Sumas de Cuadrados (SC) por su grados de libertad (gl) y se obtienen las denominadas Medias Cuadráticas (MC) y ya se puede calcular el valor del estadístico de la razón $F = MC_{\text{efecto}} / MC_{\text{error}}$ para cada una de las fuentes de varianza entre-grupos o del efecto (A, B y A x B) planteadas en el modelo de diseño del estudio.

Puede comprobarse en la Tabla 13 que existen diferencias estadísticamente significativas ($p < .05$) para el término de interacción entre los dos factores principales del modelo de diseño planteado (efecto A x B). Por tanto, el modelo que se deriva del análisis estadístico no es el de efectos aditivos sino el modelo con efectos de interacción o modelo no aditivo tal y como planteaba la hipótesis teórica del estudio.

$$\hat{Y} = M + A + B + AB$$

Tabla 3. Diseño factorial 2 x 2 entre estrés y motivación. Modelo con interacción

<i>Fuente</i>	<i>SC</i>	<i>gl</i>	<i>MC</i>	<i>F</i>	<i>p</i>	η^2
A	72.00	1	72.00	36.0	< .05	.391
B	72.00	1	72.00	36.0	< .05	.391
A x B	32.00	1	32.00	16.0	< .05	.174
Error	8.00	4	2.00			
Total	184.000	7				

El modelo que mejor se ajusta a los datos supone que la cantidad de recuerdo de las palabras dependerá no sólo del mero efecto independiente del alto o bajo

nivel de estrés y motivación; también estará relacionada con el efecto particular que cada combinación de motivación y estrés alto o bajo ocasiona a los participantes.

El cálculo del valor p de probabilidad de cada fuente de varianza se ha realizado consultando las tablas de la Razón F ($.05, 1, 4$) = 7.709. Sería necesario buscar en las tablas el valor teórico de F que corresponde a cada fuente de varianza, pero como en este diseño en los tres casos los grados de libertad son 1 para el efecto y 4 para el error pues con un valor teórico es suficiente ya que se utiliza el mismo en los tres casos. Se observa que los valores de las tres F empíricas ($F_{A1, 4} = 36, F_{B1, 4}, F_{AB1, 4} = 16$) son mayores a 7.709, luego en los tres casos se rechaza la hipótesis nula

La mejor ayuda para interpretar el efecto de interacción consiste en realizar una representación gráfica del mismo, en la que se sitúa en el eje de las abscisas los niveles de uno de los factores y en las ordenadas las puntuaciones en la variable dependiente de cada condición experimental (ver Figura 52). Se observa un efecto de interacción entre los niveles o condiciones de las variables independientes (interacción ordinal negativa). Se comprueba que los participantes que se someten a la situación de estrés bajo y con alta motivación son los que menos errores producen al recordar el listado de las palabras.

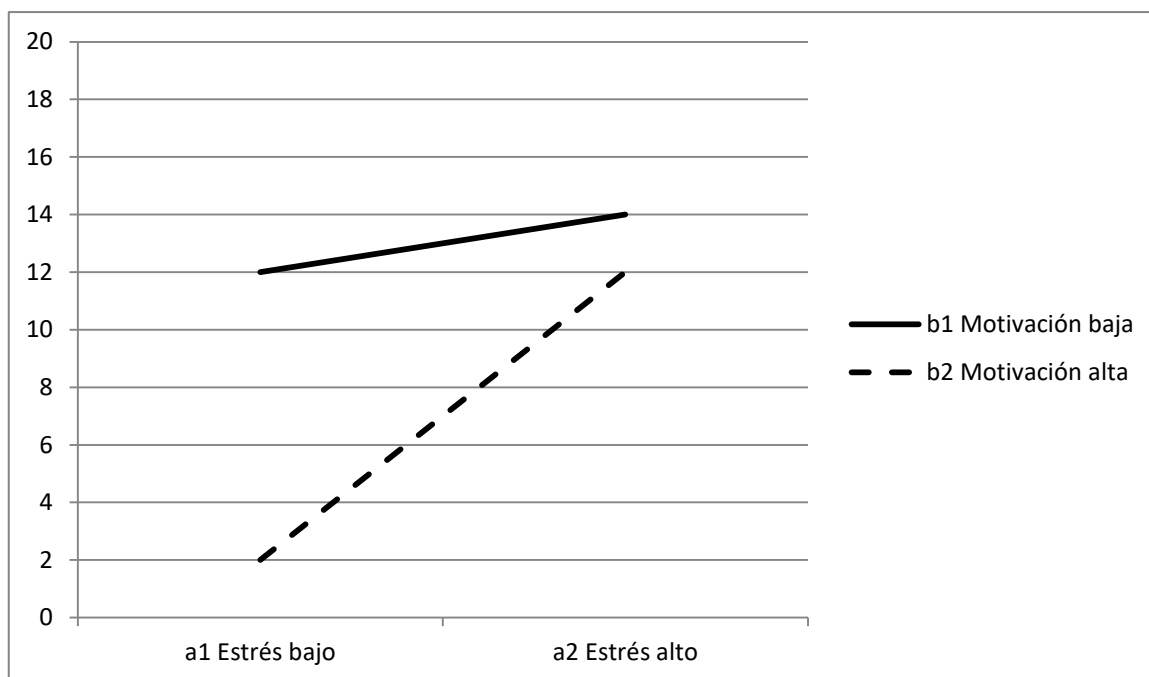


Figura 6. Representación gráfica del efecto de interacción AB

Contraste de hipótesis específicas

También en el diseño factorial, al igual que cuando se aplican diseños unifactoriales, si se rechaza la hipótesis nula hay que aplicar contrastes específicos para determinar entre qué medias se producen diferencias estadísticamente significativas, controlando la *tasa de error de Tipo I por experimento*. Las comparaciones en el diseño factorial se aplican de forma análoga al diseño unifactorial.

Como no se ha concretado ninguna hipótesis específica en el enunciado del supuesto de investigación y todas las comparaciones que se realizan incumben únicamente a las existentes entre dos grupos parece oportuno aplicar la prueba de Tukey ya que es la que tiene la máxima potencia estadística en este tipo de situaciones y controla de forma adecuada la tasa de error de tipo I.

$$|\bar{Y}_g - \bar{Y}_h| \geq \frac{q(\alpha, a, b, g_{\text{error}})}{\sqrt{2}} \sqrt{MC_{\text{error}} \sum_{i=1, j=1}^{a, b} \frac{C_{ij}^2}{n_{ij}}}$$

Se consulta la tabla de Tukey del rango estandarizado y se observa que para $q(\alpha, ab, g_{\text{error}}: q(.05, 4, 4)$ el valor del Rango Crítico de la prueba de Tukey es 5.757 (figura 53).

		Estadístico del rango estandarizado													
		Tabla Comparaciones múltiples del Test de Tukey (valores q)													
<i>gl</i>		<i>número de grupos</i>													
<i>error</i>	α	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	.05	17.97	36.98	32.82	37.08	40.41	43.12	45.40	47.36	49.07	50.6	52.0	53.2	54.3	55.4
	.01	90.03	135.0	164.3	185.6	202.2	215.8	227.2	237.0	245.6	253	260	266	272	277
2	.05	6.085	8.331	9.798	10.880	11.732	12.436	13.026	13.539	13.989	14.392	14.746	15.071	15.370	15.646
	.01	14.04	19.02	22.29	24.720	26.63	28.20	29.53	30.68	31.69	32.59	33.44	34.11	34.82	35.39
3	.05	4.501	5.910	6.825	7.502	8.037	8.478	8.853	9.177	9.462	9.714	9.941	10.149	10.340	10.516
	.01	8.261	10.62	12.17	13.33	14.24	15.00	15.64	16.20	16.69	17.13	17.51	17.86	18.19	18.51
4	.05	3.927	5.040	5.757	6.287	6.707	7.053	7.347	7.602	7.826	8.026	8.208	8.371	8.522	8.662
	.01	6.512	8.120	9.173	9.958	10.584	11.097	11.542	11.928	12.268	12.579	12.837	13.082	13.314	13.527

Figura 7. Consulta en la tabla de Tukey del valor de $q(\alpha, ab, g_{\text{error}})$

Si se ajusta la prueba de Tukey a los datos del supuesto de investigación se puede comprobar que:

$$|\bar{Y}_g - \bar{Y}_h| \geq \frac{q(0.05, 4, 4)}{\sqrt{2}} \sqrt{2 \left(\frac{1^2}{2} + \frac{-1^2}{2} + \frac{0^2}{2} + \frac{0^2}{2} \right)} \Rightarrow$$

$$\Rightarrow \frac{5.757}{\sqrt{2}} \sqrt{2} = 5.757$$

En este diseño el número total de comparaciones posibles simples dos a dos entre las medias de las condiciones es igual a seis ya que:

$$C = \frac{ab(ab - 1)}{2} = \frac{4(3)}{2} = 6$$

Se puede realizar una tabla de diferencias de medias para comprobar qué diferencia supera o iguala al valor del Rango Crítico de Tukey de 5.757 (tabla 14). Se observa que tres diferencias de medias de interacción superan dicho valor.

Tabla 4. Tabla de diferencia de medias entre las condiciones de interacción AB

Medias	a1b1 Media=12	a1b2 Media=2	a2b1 Media=14
a1b2: media 2	10	-	-
a2b1: media 14	2	12	-
a2b2: media 12	0	10	2

Analizada la distancia o la diferencia que existe entre las seis comparaciones simples, consideradas dos a dos, se puede comprobar que únicamente existen diferencias estadísticamente significativas entre el grupo con *bajo* estrés y *alta* motivación (Media de $a_1b_2 = 2$) respecto de las puntuaciones medias de las otras tres situaciones experimentales. Es decir, el valor crítico de Tukey de 5.757 únicamente se supera o se iguala cuando se compara la diferencia de medias de la condición que tiene la media de a_1b_2 (grupo de estrés bajo y motivación alta) respecto al resto de las medias de los otros tres grupos ($a_1b_2 - a_1b_1$, $a_1b_2 - a_2b_1$ y $a_1b_2 - a_2b_2$). El resto de diferencias de medias entre las medias de las condiciones no son estadísticamente significativas ($a_2b_1 - a_1b_1$, $a_2b_1 - a_1b_2$ y $a_2b_1 - a_2b_2$) porque el valor del Rango Crítico de la prueba de Tukey supera al valor de las diferencias empíricas entre las medias de los grupos o condiciones de los efectos de interacción.

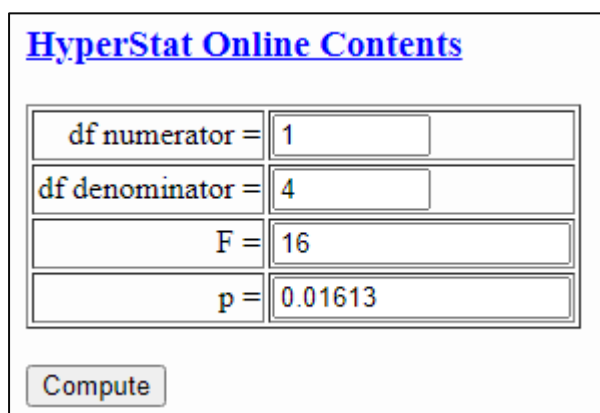
Como conclusión final de los análisis realizados, el investigador tendría que concluir que la memoria es mayor cuando los sujetos se encuentran en una situación de estrés bajo y alta motivación para recordar. Así, el efecto en la mejora de la memoria no depende del efecto del estrés y de la motivación por separado o de

forma independiente sino que el efecto de la interacción de ambas situaciones es fundamental para entender por qué la memoria es mayor.

Diseño factorial: SPSS, JASP, JAMOVI

La tabla del ANOVA del diseño factorial anterior (tabla 13) va a ser ahora completada con el valor de p exacto que les corresponde a cada fuente de varianza del tratamiento o varianza entre-grupos del modelo de diseño especificado (A, B y AB). Dicho valor de p se calcula asumiendo que la hipótesis nula es cierta (es decir, efecto o relación cero o diferencia entre las medias igual a cero) y por ello se consulta qué valor de probabilidad tendría el estadístico o resultado de $F(1, 4) = 16$ obtenido en el experimento en la distribución de la hipótesis nula que es la distribución muestral conocida del estadístico de la Razón $F(.05, 1, 4)$. Este valor p del resultado (calculado a partir de la información de la distribución muestral de la hipótesis nula) se puede obtener de forma directa con un programa estadístico como el SPSS, JASP o JAMOVI o con alguna aplicación como por ejemplo la que se ha desarrollado en la siguiente página Web: http://davidmlane.com/hyperstat/F_table.html, detallada por el profesor David Lane.

Por ejemplo, el valor de p que se corresponde con los resultados del efecto de la fuente de varianza de interacción AB ($F(1, 4) = 16$) es de .01613 (figura 54), que si se redondea en el tercer decimal sería de $p = .016$. El lector o lectora puede comprobar que el valor de p de las fuentes de varianza A ($F(1, 4) = 16$) y B ($F(1, 4) = 16$) es .00388 en los dos casos ya que tienen el mismo valor en la F empírica y los grados de libertad entre y del error también son los mismos.



HyperStat Online Contents

df numerator =	1
df denominator =	4
F =	16
p =	0.01613

Compute

Figura 8. Consulta en la tabla de Tukey del valor de $q(\alpha, ab, g_{\text{error}})$

Ahora ya se puede completar la tabla del ANOVA con los valores exactos del valor p de probabilidad que se corresponden con los resultados de las tres fuentes de varianza relacionadas con los efectos principales de las variables independientes y su interacción (tabla 13).

Tabla 5. Diseño factorial 2×2 entre estrés y motivación. Modelo con interacción

<i>Fuente</i>	<i>SC</i>	<i>gl</i>	<i>MC</i>	<i>F</i>	<i>p</i>	η^2
A	72.00	1	72.00	36.0	.004	.391
B	72.00	1	72.00	36.0	.004	.391
A x B	32.00	1	32.00	16.0	.016	.174
Error	8.00	4	2.00			
Total	184.000	7				

Otra forma de obtener el valor de p exacto de cada una de las fuentes de varianza que se plantean en el diseño de una determinada investigación es llevar a cabo el análisis con un programa estadístico. A continuación se detalla la ejecución con SPSS, JASP y JAMOVI.

SPSS

Una vez introducidos los datos en el SPSS tal y como se detalla en la figura 55, se procede al análisis acudiendo a la ventana de:

Análisis → Modelo Lineal General → Univariante

	 A estrés	 B motivación	 Yerrores
1	1	1	13
2	1	1	11
3	1	2	1
4	1	2	3
5	2	1	15
6	2	1	13
7	2	2	11
8	2	2	13
9			

Figura 9. Base de datos de un diseño factorial A x B con SPSS

El proceso de acceso al diseño elegido en el SPSS es entonces el especificado en la figura 56.

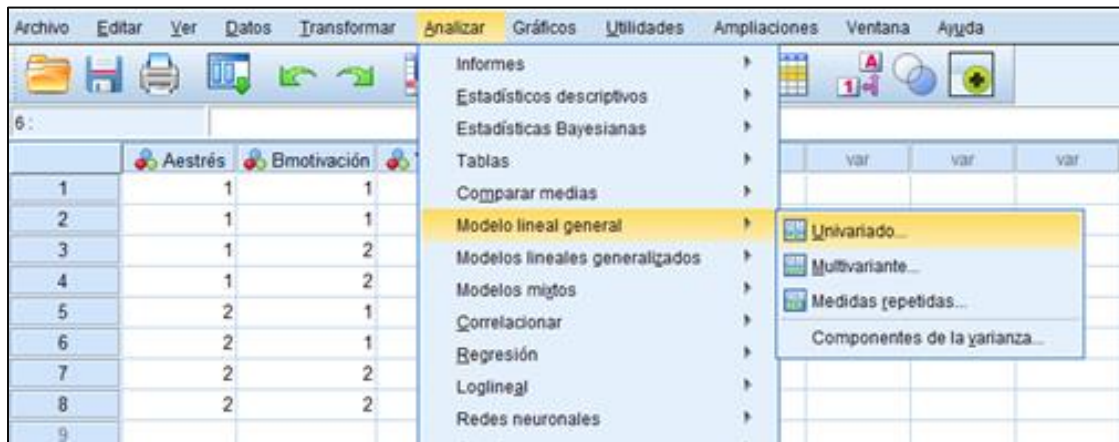
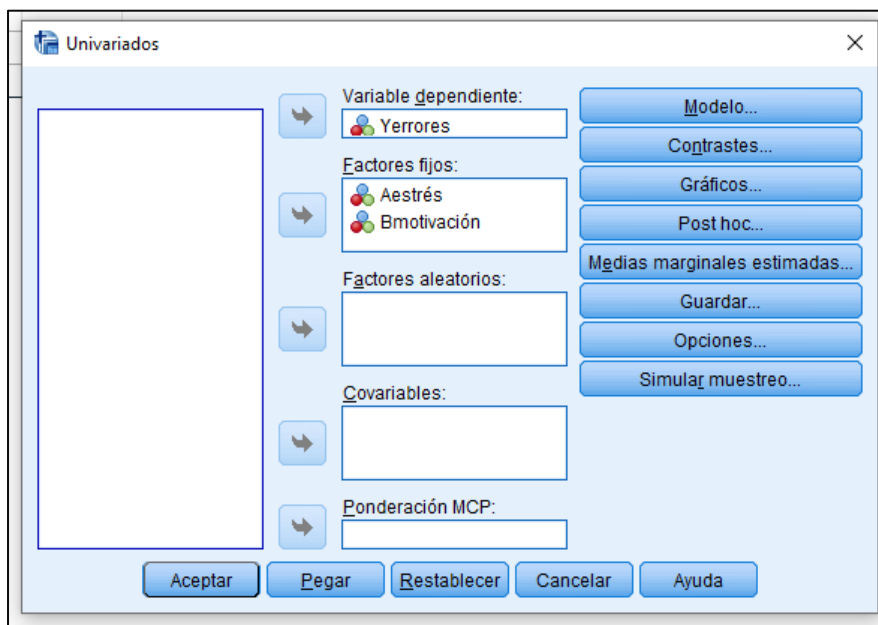


Figura 10. Acceso al diseño univariado con el SPSS

A continuación se sitúa cada una de las variables que se desea estudiar en el lugar correspondiente para llevar a cabo el análisis y en Opciones se seleccionan estadísticos y estimaciones del tamaño del efecto y también se selecciona el apartado de Medias marginales estimadas para que el output o salida de resultados contenga las medias marginales de los efectos principales A y B y las medias de las celdillas de interacción AB (figura 57).



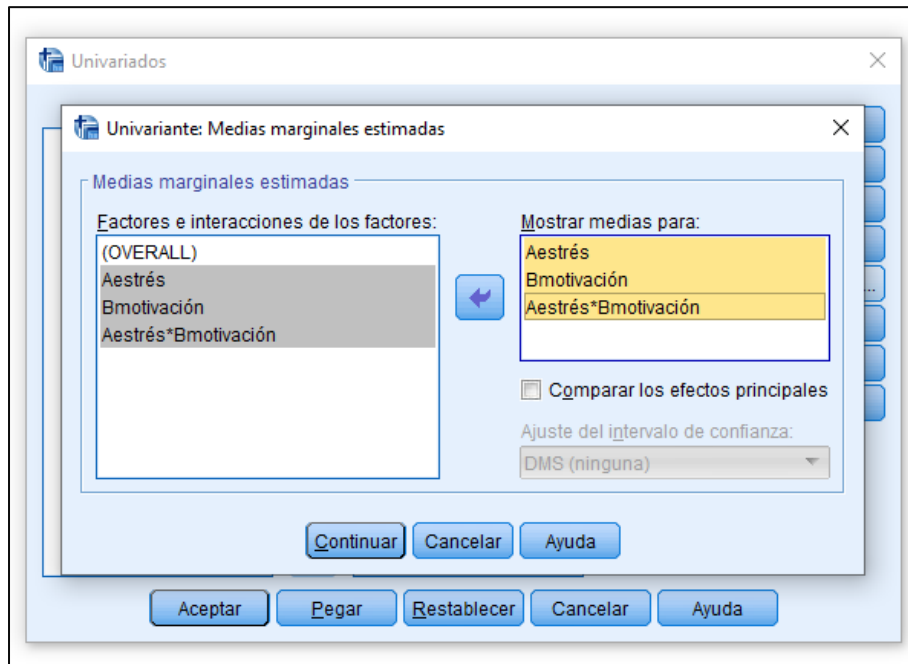


Figura 11. Acomodación de las variables junto con opciones complementarias de análisis y pantalla de las medias marginales estimadas con SPSS

El SPSS ofrece los siguientes resultados. En primer lugar, detalla el número de observaciones que hay en cada condición de los efectos principales A y B (Factores inter-sujetos). En segundo lugar, incluye los estadísticos descriptivos de la fuente de varianza de la interacción: media desviación típica y número de observaciones (Estadísticos descriptivos). En tercer lugar, se presenta la tabla del ANOVA que incluye el valor de la proporción de varianza explicada (η_p^2) ya que se le solicitó en Opciones junto con los estadísticos descriptivos anteriores (figura 58). El lector o lectora puede comprobar que el programa SPSS ofrece los resultados de la eta cuadrado parcial y por ello no coincide con los valores de η^2 que se han calculado en las tablas de ANOVA anteriores. Para una explicación detallada se puede consultar el apartado dedicado al tamaño del efecto que se ha presentado anteriormente. Se recomienda utilizar los valores de η^2 para elaborar el informe de resultados ya que la η_p^2 sobreestima el valor del tamaño del efecto.

Factores inter-sujetos		
		N
Aestrés	1	4
	2	4
Bmotivación	1	4
	2	4

Estadísticos descriptivos				
Variable dependiente: Yerrores				
Aestrés	Bmotivación	Media	Desv. Desviación	N
1	1	12,00	1,414	2
	2	2,00	1,414	2
	Total	7,00	5,888	4
2	1	14,00	1,414	2
	2	12,00	1,414	2
	Total	13,00	1,633	4
Total	1	13,00	1,633	4
	2	7,00	5,888	4
	Total	10,00	5,127	8

Pruebas de efectos inter-sujetos						
Variable dependiente: Yerrores						
Origen	Tipo III de suma de cuadrados	gl	Media cuadrática	F	Sig.	Eta parcial al cuadrado
Modelo corregido	176,000 ^a	3	58,667	29,333	,003	,957
Intersección	800,000	1	800,000	400,000	,000	,990
Aestrés	72,000	1	72,000	36,000	,004	,900
Bmotivación	72,000	1	72,000	36,000	,004	,900
Aestrés * Bmotivación	32,000	1	32,000	16,000	,016	,800
Error	8,000	4	2,000			
Total	984,000	8				
Total corregido	184,000	7				

a. R al cuadrado = ,957 (R al cuadrado ajustada = ,924)

Figura 12. Resultados del ANOVA entre-grupos A x B con SPSS

En quinto lugar, el programa SPSS detalla las medias marginales de los efectos principales A y B y de la interacción ya que se le solicitó al programa SPSS cuando se seleccionó el apartado de Medias marginales estimadas (figura 59). El valor de 1 indica 'bajo' y el valor de 2 indica 'alto' en la condición del factor o variable independiente.

Medias marginales estimadas

1. Aestrés

Variable dependiente: Yerrores

Aestrés	Media	Desv. Error	Intervalo de confianza al 95%	
			Límite inferior	Límite superior
1	7,000	,707	5,037	8,963
2	13,000	,707	11,037	14,963

2. Bmotivación

Variable dependiente: Yerrores

Bmotivación	Media	Desv. Error	Intervalo de confianza al 95%	
			Límite inferior	Límite superior
1	13,000	,707	11,037	14,963
2	7,000	,707	5,037	8,963

3. Aestrés * Bmotivación

Variable dependiente: Yerrores

Aestrés	Bmotivación	Media	Desv. Error	Intervalo de confianza al 95%	
				Límite inferior	Límite superior
1	1	12,000	1,000	9,224	14,776
	2	2,000	1,000	-,776	4,776
2	1	14,000	1,000	11,224	16,776
	2	12,000	1,000	9,224	14,776

Figura 13. Resultados del ANOVA entre-grupos A x B con SPSS

El programa SPSS no calcula la prueba de Tukey (u otras pruebas de hipótesis específicas) para el efecto de la interacción de dos o más factores. Solamente calcula esas pruebas de contraste de hipótesis específicas para los efectos principales del modelo de diseño. Por lo tanto, si se utiliza en SPSS para analizar un diseño factorial y se desea conocer qué pares de medias simples del efecto de interacción difieren entre sí de forma estadísticamente significativa será necesario calcular el Rango Crítico, teórico o tabular de Tukey y después comparar si la diferencia de medias empírica obtenida en el estudio supera o iguala dicha

diferencia para poder concluir que esa diferencia tiene un valor de p menor o igual al valor del alfa. Si es así, se puede concluir que esa diferencia entre las dos medias es estadísticamente significativa.

Otra solución para continuar con el estudio del efecto de interacción es introducir los datos como si fuese un diseño unifactorial con cuatro condiciones ($A = 4$), utilizando cada efecto de interacción como si fuese una condición del diseño (por ejemplo, A_1, B_1 se corresponde con la condición A_1 del diseño unifactorial $A = 4$) y llevar a cabo un diseño unifactorial solicitando la prueba de comparaciones múltiples en el programa SPSS. En la figura siguiente (figura 60) se puede observar la introducción de datos para estudiar el efecto de interacción en un diseño AB con el programa SPSS. Se ejecuta ese diseño unifactorial y directamente se observan los resultados de la prueba de hipótesis específicas (por ejemplo, la prueba HSD de Tukey), sin mirar el ANOVA que carece de sentido para el supuesto de investigación.

	Medida	GrupoVI
1	13	1
2	11	1
3	1	2
4	3	2
5	15	3
6	13	3
7	11	4
8	13	4

Con los valores de las etiquetas:

	Medida	GrupoVI
1	13	A1B1
2	11	A1B1
3	1	A1B2
4	3	A1B2
5	15	A2B1
6	13	A2B1
7	11	A2B2
8	13	A2B2

Figura 14. Base de un diseño $A = 4$ cuyo objetivo es estudiar el efecto de interacción de un diseño 2×2 con el SPSS

Se ejecuta el diseño unifactorial $A = 4$ y en Opciones se solicita por ejemplo la prueba de Tukey para el factor o variable independiente de GrupoVI que tiene cuatro condiciones. Los resultados se detallan en la Figura 61. Se observa que las diferencias entre los pares de medias que son estadísticamente significativas se encuentran entre a_1b_1 respecto a las condiciones de a_1b_2 , a_2b_1 y a_2b_2 . Son las mismas diferencias que las que se detectan con el uso del Rango Crítico de la prueba de Tukey.

Pruebas post hoc

GrupoVI

Comparaciones múltiples

Variable dependiente: Medida
HSD Tukey

(I) GrupoVI	(J) GrupoVI	Diferencia de medias (I-J)	Desv. Error	Sig.	Intervalo de confianza al 95%	
					Límite inferior	Límite superior
A1B1	A1B2	10,00*	1,414	,007	4,24	15,76
	A2B1	-2,00	1,414	,553	-7,76	3,76
	A2B2	,00	1,414	1,000	-5,76	5,76
A1B2	A1B1	-10,00*	1,414	,007	-15,76	-4,24
	A2B1	-12,00*	1,414	,004	-17,76	-6,24
	A2B2	-10,00*	1,414	,007	-15,76	-4,24
A2B1	A1B1	2,00	1,414	,553	-3,76	7,76
	A1B2	12,00*	1,414	,004	6,24	17,76
	A2B2	2,00	1,414	,553	-3,76	7,76
A2B2	A1B1	,00	1,414	1,000	-5,76	5,76
	A1B2	10,00*	1,414	,007	4,24	15,76
	A2B1	-2,00	1,414	,553	-7,76	3,76

Se basa en las medias observadas.
El término de error es la media cuadrática(Error) = 2,000.
*. La diferencia de medias es significativa en el nivel ,05.

Figura 15. Resultados de la prueba HSD de Tukey con el diseño A = 4 cuyo objetivo es estudiar el efecto de interacción de un diseño 2 x 2

JASP

Desde el programa gratuito JASP se puede llamar a la base de datos que se haya creado con Excel o con SPSS, por ejemplo, y ejecutar posteriormente el diseño entre-sujetos univariado 2 x 2 (figura 62).

The image shows two parts of the JASP interface. On the left is a data table with 8 rows and 4 columns. The columns are labeled 'Aestrés', 'Bmotivación', and 'Errores'. The rows contain numerical values for each cell. On the right is the ANOVA configuration window. The 'Dependent Variable' is set to 'Errores'. The 'Fixed Factors' are 'Aestrés' and 'Bmotivación'. Under the 'Display' section, several options are checked: 'Descriptive statistics', 'Estimates of effect size', 'eta squared', 'partial eta squared', and 'omega squared'.

	Aestrés	Bmotivación	Errores
1	1	1	13
2	1	1	11
3	1	2	1
4	1	2	3
5	2	1	15
6	2	1	13
7	2	2	11
8	2	2	13

Figura 16. Base de datos y comodación de las variables junto con opciones complementarias de análisis con JASP

En el programa JASP se pueden solicitar más estadísticos del tamaño del efecto. Concretamente, se ha solicitado que se calculen los valores de η^2 y η_p^2 y también el de omega cuadrado (ω^2). El investigador o investigadora debe interpretar el que considere más conveniente para detallar en su informe de investigación.

Los resultados que ofrece el programa JASP se detallan en la figura 63.

Results									
ANOVA									
ANOVA - Yerrores									
Cases	Sum of Squares	df	Mean Square	F	p	η^2	η_p^2	ω^2	
Aestrés	72.00	1	72.00	36.00	3.88e-3	0.39	0.90	0.38	
Bmotivación	72.00	1	72.00	36.00	3.88e-3	0.39	0.90	0.38	
Aestrés * Bmotivación	32.00	1	32.00	16.00	0.02	0.17	0.80	0.16	
Residuals	8.00	4	2.00						

Note. Type III Sum of Squares

Descriptives				
Descriptives - Yerrores				
Aestrés	Bmotivación	Mean	SD	N
1	1	12.00	1.41	2
	2	2.00	1.41	2
2	1	14.00	1.41	2
	2	12.00	1.41	2

Figura 17. Resultados del ANOVA entre-grupos A x B con JASP

En el apartado que tiene el programa JASP de medias marginales se puede solicitar que calcule las medias marginales de los efectos principales A y B y las medias de los efectos de interacción AB (figura 64).

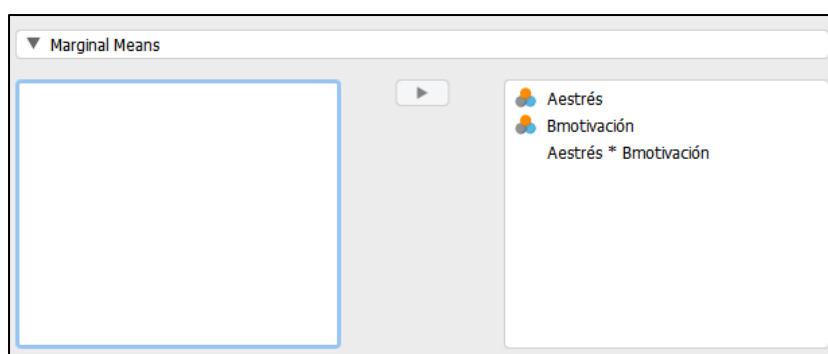


Figura 18. Seleccionar el análisis de las medias marginales con JASP

El resultado que detalla el programa JASP de las medias marginales de A, B y AB se puede observar en la figura 65. El programa JASP ofrece las medias y la

error estándar (SE, Standard Error) que no hay que confundir con la desviación típica. Como ofrece el intervalo de las medias detalla el SE ya que el intervalo se construye con el error estándar. Cuando se realice la redacción de los resultados se utilizarán los descriptivos de medias (M), desviación típica (DT) y número de observaciones (N).

Marginal Means					
Marginal Means - Aestrés					
Aestrés	Marginal Mean	95% CI for Mean Difference		SE	
		Lower	Upper		
1	7.00	5.04	8.96	0.71	
2	13.00	11.04	14.96	0.71	
Marginal Means - Bmotivación					
Bmotivación	Marginal Mean	95% CI for Mean Difference		SE	
		Lower	Upper		
1	13.00	11.04	14.96	0.71	
2	7.00	5.04	8.96	0.71	
Marginal Means - Aestrés * Bmotivación					
Aestrés	Bmotivación	Marginal Mean	95% CI for Mean Difference		SE
			Lower	Upper	
1	1	12.00	9.22	14.78	1.00
2	1	14.00	11.22	16.78	1.00
1	2	2.00	-0.78	4.78	1.00
2	2	12.00	9.22	14.78	1.00

Figura 19. Medias marginales con JASP

Una de las ventajas del programa JASp es que sí realiza un análisis de hipótesis específicas para el efecto de interacción AB del diseño factorial. Para ejecutar los estadísticos de hipótesis específicas se selecciona Post Hoc Tests y se activa la fuente de varianza de la interacción (Aestrés * Bmotivación), por ejemplo. El programa JASP ofrece directamente las 6 diferencias de medias que hay que analizar y no 12 como el SPSS. Conviene recordar que si se utiliza el SPSS hay que tener en cuenta que dichas diferencias de medias están duplicadas, pues solo hay que interpretar seis diferencias simples entre los pares de medias.

En la figura 66 se detalla cómo se solita la prueba post hoco para el efecto de interacción AB.

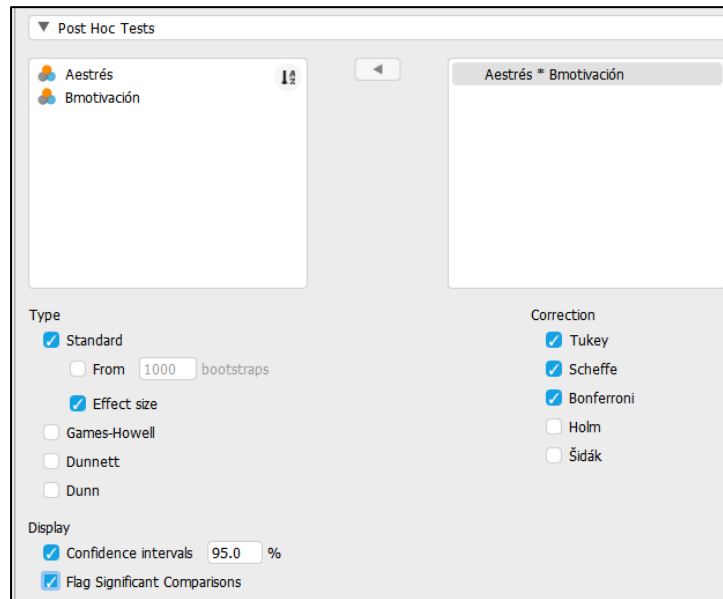


Figura 20. Selección del contraste de hipótesis específicas con JASP

El programa JASP ofrece un conjunto de resultados sobre las diferencias de medias entre todos los pares de medias y acaba con las columnas de los estadísticos de contraste de hipótesis específicas (figura 67). El investigador o investigadora debe seleccionar aquella prueba que sea la más adecuada para el modelo de diseño que está utilizando en su investigación, teniendo en cuenta que debe controlar de forma adecuada la tasa de error de Tipo I y debe ser el más potente (el que tenga menor error de Tipo II) para maximizar la validez de conclusión estadística. Un inconveniente del programa JASP es que no estima el valor del tamaño del efecto (d de Cohen) para cada diferencia entre un par de medias simples, siendo necesario recurrir a otro programa para calcularlo como la página Web de la Colaboration Campbell, o se puede utilizar el programa JAMOVl que sí calcula dicho valor de tamaño del efecto con el estadístico de diferencia estandarizada de medias d de Cohen.

Post Hoc Tests

Standard

Post Hoc Comparisons - Aestrés * Bmotivación

		95% CI for Mean Difference			SE	t	P _{Tukey}	P _{Scheffe}	P _{bonf}
		Mean Difference	Lower	Upper					
1, 1	2, 1	-2.00	-7.76	3.76	1.41	-1.41	0.55	0.62	1.00
	1, 2	10.00	4.24	15.76	1.41	7.07	7.28e-3**	0.01*	0.01*
	2, 2	-4.44e-15	-5.76	5.76	1.41	-3.14e-15	1.00	1.00	1.00
2, 1	1, 2	12.00	6.24	17.76	1.41	8.49	3.68e-3**	5.10e-3**	6.35e-3**
	2, 2	2.00	-3.76	7.76	1.41	1.41	0.55	0.62	1.00
1, 2	2, 2	-10.00	-15.76	-4.24	1.41	-7.07	7.28e-3**	0.01*	0.01*

* p < .05, ** p < .01
 Note. P-value and confidence intervals adjusted for comparing a family of 4 estimates (confidence intervals corrected using the tukey method).

Figura 21. Resultados del contraste de hipótesis específicas con JASP

JAMOVI

El programa JAMOVI puede abrir los ficheros de SPSS, JASP o Excel, por ejemplo y la presentación de su base de datos es muy semejante (figura 68).

The image shows the JAMOVI interface. On the left is a data table with columns 'Aestrés', 'Bmotivación', and 'Errores'. On the right is the 'ANOVA' configuration window where 'Errores' is set as the dependent variable and 'Aestrés' and 'Bmotivación' are set as fixed factors. The 'Effect Size' section is checked for η^2 , partial η^2 , and ω^2 .

Figura 22. Base de datos y comodación de las variables junto con opciones complementarias de análisis con JAMOVI

Los resultados con el programa JAMOVI del ANOVA junto con los estadísticos del tamaño del efecto solicitados se detallan en la figura 69.

ANOVA

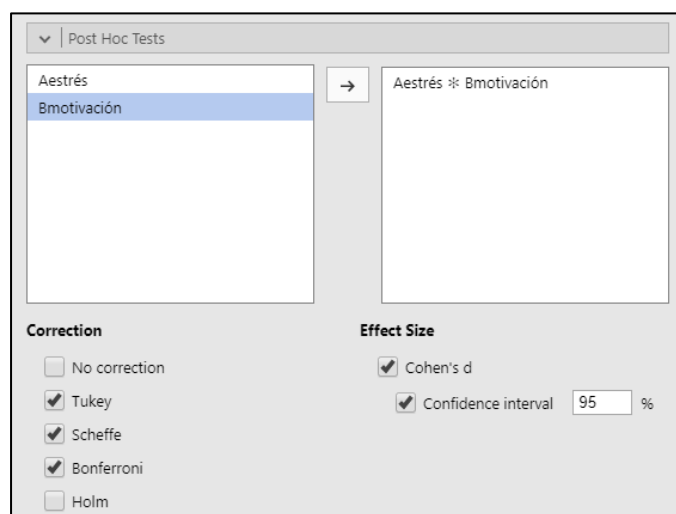
ANOVA - Errores

	Sum of Squares	df	Mean Square	F	p	η^2	η^2p	ω^2
Aestrés	72.00	1	72.00	36.0	0.004	0.391	0.900	0.376
Bmotivación	72.00	1	72.00	36.0	0.004	0.391	0.900	0.376
Aestrés * Bmotivación	32.00	1	32.00	16.0	0.016	0.174	0.800	0.161
Residuals	8.00	4	2.00					

[3]

Figura 23. Resultados del ANOVA entre-grupos A x B con JAMOVI

Como se observa en la figura 70, el programa JAMOVl puede efectuar las pruebas post o estadísticos de contraste de hipótesis específicas y, además, ofrece para cada diferencia entre un par de medias el valor del tamaño del efecto d de Cohen denominado diferencia estandarizada de medias. Por lo tanto, de los tres programas estadísticos analizados, solamente el programa JAMOVl calcula los valores del tamaño del efecto para el efecto de interacción AB, es decir, calcula una diferencia de medias estandarizada d de Cohen para cada uno de los pares de medias que tiene el diseño del estudio junto con su intervalo de confianza. Conviene tener presente que esos valores tan altos de d de Cohen son irreales en la investigación real y son propios de un diseño cuyos datos se han programado para la docencia. Siguiendo la propuesta de Jacob Cohen, un tamaño del efecto de diferencia estandarizada de media de $d = 0.2$ es pequeño, $d = 0.5$ es medio y un valor de $d = 0.8$ sería grande. Y, gran parte de las terapias psicológicas tienen en la realidad un tamaño del efecto en torno a $d = 0.5$, es decir, un tamaño del efecto mediano.



Post Hoc Tests

Post Hoc Comparisons - Aestres * Bmotivación

Comparison		Mean Difference	SE	df	t	Ptukey	Pscheffe	Pbonferroni	Cohen's d	95% Confidence Interval		
Aestres	Bmotivación									Aestres	Bmotivación	Lower
1	1	- 1 2	10.00	1.41	4.00	7.07	0.007	0.010	0.013	7.07	-0.405	14.547
		- 2 1	-2.00	1.41	4.00	-1.41	0.553	0.615	1.000	-1.41	-4.518	1.690
	- 2 2	-4.44e-15	1.41	4.00	-3.14e-15	1.000	1.000	1.000	0.00	-2.776	2.776	
2	2	- 2 1	-12.00	1.41	4.00	-8.49	0.004	0.005	0.006	8.49	-0.295	17.265
		- 2 2	-10.00	1.41	4.00	-7.07	0.007	0.010	0.013	-7.07	-14.547	0.405
	1	- 2 2	2.00	1.41	4.00	1.41	0.553	0.615	1.000	1.41	-1.690	4.518

Note. Comparisons are based on estimated marginal means

Figura 24. Resultados de las pruebas de hipótesis específicas para el efecto de interacción AB y valores de d de Cohen con JAMOVl

Estimar el tamaño del efecto con la Colaboración Campbell

Se continuara en el documento de apuntes de factorial_2

Redacción de los resultados de un diseño entre-sujetos A x B

Se continuara en el documento de apuntes de factorial_2

Diseño entre-sujetos factorial: más de dos factores

Se continuara en el documento de apuntes de factorial_2