

CAPÍTULO

5

DISEÑOS EXPERIMENTALES: SEGUNDA PARTE

INTRODUCCIÓN

El presente capítulo amplía algunos de los conceptos vertidos en el capítulo 7 de *Metodología de la investigación*, 5ª edición, parte impresa; en particular en lo referente a las fuentes de invalidación interna, la equivalencia de los grupos experimentales y la técnica de emparejamiento. Asimismo, se revisan otros diseños experimentales: las series cronológicas múltiples y los factoriales. Finalmente, se comentan los diseños cuasiexperimentales.

FUENTES DE INVALIDACIÓN INTERNA (AMPLIACIÓN Y EJEMPLIFICACIÓN)

En el libro impreso se describieron estas fuentes y se mencionó que se deben eliminar mediante el control, lo que León y Montero (2003) denominan: “descartar causas alternativas”. A continuación se amplían y ejemplifican.

1. *Historia*. Son acontecimientos que ocurren durante el desarrollo del experimento, que pueden afectar a la variable dependiente y llegan a confundir los resultados experimentales. Por ejemplo, durante un experimento para analizar el efecto que tienen distintas formas de realimentación en el trabajo en la motivación (vía el supervisor “cara a cara” y por escrito); considerando dos

grupos de obreros, a uno le aumentan el salario o se le felicita por su desempeño en el trabajo mientras está llevándose a cabo el experimento y al otro grupo no. Diferencias en la variable dependiente pueden atribuirse a la manipulación de la independiente o al acontecimiento que ocurrió durante el experimento.

2. *Maduración.* Son procesos internos de los participantes que operan como consecuencia del tiempo y afectan los resultados del experimento, tales como cansancio, hambre, aburrimiento, aumento en la edad y cuestiones similares. Si a un grupo de niños se le expone a un nuevo método de aprendizaje por dos años, los resultados pueden estar influidos simplemente por la maduración de los infantes durante el tiempo que persistió el experimento. En un experimento quizá los sujetos se cansen y sus respuestas sean afectadas por ello. Si tenemos dos grupos y la condición experimental del primero implica mucho más tiempo que la del segundo, puede afectar esta fuente.
3. *Inestabilidad del instrumento de medición.* Poca o nula confiabilidad de las mediciones, que provoca desconfianza en las puntuaciones obtenidas, no sabemos si las deducciones son certeras, el instrumento a veces da unos resultados y luego, otros. Esto se evita diseñando instrumentos válidos y confiables (vea el capítulo 9 de la obra impresa: “Recolección de los datos cuantitativos”).
4. *Inestabilidad del ambiente experimental.* Tal como se mencionó en el capítulo 7 del texto impreso, las condiciones o componentes del ambiente o entorno del experimento pueden ser desiguales para todos los grupos participantes. Imaginemos que en un experimento sobre memorización se tienen dos grupos, y al sortearlos a uno le corresponde realizarlo en un aula cerca de donde se efectúa una remodelación del edificio y al otro, en un aula lejos de tal remodelación. Si el experimento requiriera de elevada concentración por parte de los individuos que participan en él, pudiera ser que la concentración de un grupo fuera diferente de la del otro (a causa del ruido, del tránsito de personas, etc.), y ello afectara la interpretación de los resultados. Tal vez las diferencias en los grupos se deban a variaciones en la concentración y no a la variable independiente, o es posible que ésta sí tenga un efecto; pero no podemos estar seguros de ello. No tenemos confianza en los resultados.
5. *Administración de pruebas.* Se refiere al efecto que puede tener la aplicación de una prueba sobre las puntuaciones de pruebas subsecuentes. Por ejemplo, si en un experimento sobre prejuicio en el cual tenemos un grupo al que se le aplica una prueba para detectar su nivel de prejuicio, luego se le expone a un estímulo

experimental (que supuestamente debe reducir el prejuicio), y después se vuelve a medir el prejuicio para evaluar si disminuyó o no. Puede ocurrir que las puntuaciones de prejuicio de la segunda medición (después del estímulo) sean influidas por la aplicación de la primera prueba sobre prejuicio (antes del estímulo). Es decir, la administración de la primera medición puede sensibilizar a los participantes del experimento y, cuando respondan a la segunda prueba, sus respuestas estarían afectadas por esa sensibilización. Si disminuye el prejuicio y no hay control, no sabremos cuánto se debió al estímulo experimental o a la variable independiente, y cuánto a dicha sensibilización.

6. *Instrumentación.* Hace referencia a cambios en los instrumentos de medición o en los observadores, los cuales son capaces de producir variaciones en los resultados que se obtengan. Si la prueba del grupo experimental es diferente a la del grupo de control u otro grupo experimental, puede intervenir la instrumentación. Por ejemplo, si queremos ver el efecto de dos diferentes métodos de enseñanza, a un grupo lo exponemos a un método, mientras que a otro grupo al otro método; y después aplicamos un examen de conocimientos para ver la eficacia de cada método y comparar los resultados. Si los exámenes no fueran equivalentes podría presentarse esta fuente. Imaginemos que el examen de un grupo es más fácil que el examen del otro grupo, ¿cómo estaremos seguros de que las diferencias en las puntuaciones de los exámenes se deben al estímulo (método de enseñanza) y no a que se trata de mediciones distintas?
7. *Regresión estadística.* Se refiere a que a veces seleccionamos participantes sobre la base de puntuaciones extremas y cuando se miden por primera vez se encuentran en valores muy altos o bajos en la variable que nos interesa, después tienden a regresar a su estado normal, y en una segunda medición obtienen valores no extremos; la comparación entre las dos mediciones indica un cambio, pero en realidad lo que ocurre es que los medimos cuando su condición en la variable era extrema.

En términos de Campbell y Stanley (1966, pp. 24-28): “se trata de un efecto provocado por una tendencia que muestran los y las participantes seleccionados sobre la base de puntuaciones extremas, al regresar en pruebas posteriores, a un promedio en la variable en la que fueron elegidos”.

Por ejemplo, si pretendemos evaluar el efecto del liderazgo autocrático del profesor en la ansiedad de los alumnos y primero aplicáramos al grupo una prueba de ansiedad, para posteriormente exponerlo a un profesor autocrático y

volver a medir su ansiedad. Y entonces los participantes se encontraran bastante ansiosos durante la aplicación de la primera prueba (porque tienen un examen difícil al día siguiente); podría ocurrir que se observara que están “menos ansiosos” después de recibir la manipulación, es decir, aparentemente el profesor autocrático redujo la ansiedad. Pero en realidad lo que sucedió fue que durante la segunda prueba estaban “retornando” a su nivel común de ansiedad.

La regresión estadística representa el hecho de que puntuaciones extremas en una distribución particular tenderán a desplazarse (esto es, regresar) hacia el promedio de la distribución en función de mediciones repetidas (Christensen, 2006). Entre una primera y una segunda mediciones, las puntuaciones más altas tienden a bajar y las más bajas a aumentar. En ocasiones, este fenómeno de regresión se presenta porque ambas mediciones no están adecuadamente correlacionadas.

8. *Selección.* Puede presentarse al elegir a las personas para los grupos del experimento, de tal manera que los grupos no sean equiparables. Es decir, si no se escogen los participantes de los grupos asegurándose su equivalencia, la selección resultaría tendenciosa. Por ejemplo, en un experimento sobre métodos educativos, si en un grupo se incluyen a los estudiantes más inteligentes y estudiosos; y en otro grupo a los estudiantes menos inteligentes y estudiosos, las diferencias entre los grupos se deberán a una selección tendenciosa, aparte del tratamiento experimental o variable independiente.
9. *Mortalidad experimental.* Se refiere a diferencias en la pérdida de participantes entre los grupos que se comparan. Si en un grupo se pierde 25% de los participantes y en otro grupo sólo 2%, los resultados se verían influidos por ello, además de por el tratamiento experimental. La pérdida de participantes puede deberse a diversas razones, una de ellas es la desmoralización (Grinnell, Williams y Unrau, 2009), por lo que algunos autores la incluyen como una fuente de invalidación interna por sí misma. Imaginemos un experimento que utiliza como estímulo un programa de televisión antisocial que ha sido visto por una tercera parte del grupo al que se le expondrá, mientras que al otro grupo se le expone a un programa prosocial que nadie ha visto. Condiciones agresivas, dolorosas, de cansancio, etc., pueden provocar mortalidad diferencial en los grupos, y ésta puede ocurrir no sólo por el experimento en sí, sino por el tipo de personas que componen cada grupo o por factores externos al experimento.

- 10. *Interacción entre selección y maduración.*** Se trata de un efecto de maduración que no es igual en los grupos del experimento, debido a algún factor de selección. Ésta puede dar origen a diferentes tasas de maduración o cambio autónomo entre grupos. Por ejemplo, si seleccionamos para un grupo a participantes que acostumbran alimentarse a cierta hora (12:00 p.m.) y para el otro, a personas que se alimentan a otra hora (3:00 p.m.), y el experimento se lleva a cabo de 11:00 a.m. a 2:30 p.m., la selección tendenciosa puede provocar un efecto de maduración distinto en los dos grupos: hambre. Esto ocurriría si en una convención de una empresa multinacional experimentamos con mujeres ejecutivas de distintos países latinoamericanos (cuyo almuerzo o comida es a horas muy distintas), y no tomamos en cuenta la interacción que llegue a darse entre la selección y la maduración.
- 11. *Difusión de tratamientos experimentales.*** Se refiere a que los participantes de los grupos experimentales y de control intercambien entre sí información sobre la naturaleza del experimento, en particular respecto al estímulo, lo cual podría “nublar” los efectos de éste. Por ejemplo, si un grupo recibe dinero por participar y el otro no, si los integrantes de este último grupo se enteran de que a sus compañeros les están pagando, podrían protestar y desconcertar a toda la muestra del estudio, lo cual afectaría al experimento. Si en el caso de Naves y Poplawsky (1984), los estudiantes hubieran comentado a los demás participantes que se les “aparecía” un deficiente mental y esta información se difunde (se “corre la voz”), los siguientes participantes no actuarían con naturalidad y el experimento resultaría un fracaso. Esta fuente resulta un riesgo cuando los grupos del experimento mantienen cercanía o contacto, pero se puede evitar eligiendo participantes para la muestra que pertenezcan a diferentes segmentos (en el ejemplo, estudiantes de diferentes carreras y semestres). Cuando se sospeche la presencia de la difusión de tratamientos, conviene entrevistar a los participantes para evaluar el grado de comunicación entre grupos y sus consecuencias.
- 12. *Actuaciones anormales del grupo de control.*** Consiste en que si el grupo testigo conoce su condición, se esfuerce en un grado superlativo con el fin de obtener puntuaciones más favorables en la variable dependiente. Por ejemplo, en un experimento para probar un nuevo método educativo, los participantes del grupo que no aprenden por medio de este método innovador, al deducir que se

espera que ellos o ellas obtengan las puntuaciones más bajas, se esfuercen por encima de lo normal y sus calificaciones sean elevadas.

13. *Compensación.* Los beneficios de un experimento pueden ser desiguales o inequitativos entre los grupos, entonces podría ocurrir que el grupo de control lo resintiera y esto afecte los resultados (Creswell, 2009). Por ejemplo, si el grupo experimental recibiera una terapia y el de control no, los individuos de este último pueden sentirse devaluados o desmoralizados; esto es particularmente crítico cuando a un grupo se le otorgan estímulos económicos, como ya se señaló.
14. *Otras interacciones.* Es posible que haya diversos efectos provocados por la interacción de las fuentes de invalidación interna. La selección pudiera interactuar con la mortalidad experimental, la historia con la maduración, la maduración con la inestabilidad, etc. También pueden afectar simultáneamente varias de estas fuentes y la validez interna se deteriora aún más. Cada vez que hay presencia de estas fuentes, será mayor nuestra incertidumbre con respecto a las causas que produjeron cambios en las variables dependientes. Si no hubo cambios, no sabemos si se debió a que una o más fuentes contrarrestaron los posibles efectos del tratamiento. Por ejemplo, el método de enseñanza más exitoso se utilizó con los participantes menos motivados y el menos exitoso con los más motivados, lo cual provocó que se compensaran ambos factores. De este modo, no sabremos cómo interpretar los resultados.

Los sujetos participantes y el experimentador como fuentes de invalidación interna

Otra razón que llega a atentar contra la interpretación correcta y certera de los resultados de un experimento es el comportamiento de los propios participantes. Es posible que ellos ingresen al experimento con ciertas actitudes, expectativas y prejuicios que alteren su comportamiento durante el estudio. Por ejemplo, no colaborar y actuar negativamente, hasta el punto de llegar a ser hostiles. Lo anterior debe tenerse presente antes y durante la investigación. Es necesario evaluar qué participantes pueden arruinar el experimento y descartarlos, o procurar que en todos los grupos haya personas con actitudes positivas y negativas (si quienes tienen actitudes negativas van a un único grupo, la validez interna estaría en problemas). Recordemos que las personas que intervienen en un experimento, de una manera u otra, tienen motivos precisamente para esa participación y su papel será activo en muchas ocasiones.

Además, el mismo experimentador o experimentadora pueden afectar los resultados de la investigación. El experimentador no es un observador pasivo que no interactúa, sino un observador activo que llega a influir en los resultados del estudio (Christensen, 2006). El experimentador tiene una serie de motivos que lo llevan a realizar su experimento y con él desea probar sus hipótesis. Ello, consciente o inconscientemente, puede conducir a que afecte el comportamiento de los sujetos en dirección de su hipótesis. Por ejemplo, dar explicaciones más completas a uno de los grupos. Hay que evitar lo anterior, y en varios casos quien trate con los participantes no debe ser el experimentador, sino alguien que no conozca la hipótesis, las condiciones experimentales ni los propósitos del estudio, sino que sólo reciba instrucciones precisas sobre lo que debe hacer y cómo desempeñarse.

Los participantes en el experimento tampoco deben conocer las hipótesis ni las condiciones experimentales; incluso, con frecuencia es necesario distraerlos de los verdaderos propósitos del estudio, aunque al finalizar se les deba dar una explicación completa del experimento. Cuando, por ejemplo, se analizan los efectos de medicamentos, los investigadores hacen creer a un grupo que se le está administrando medicamentos cuando en realidad no es así, sino que se le da píldoras de azúcar. Esto evita la influencia que la expectativa de recibir medicamento pudiera tener en la variable dependiente. A esta sustancia que no tiene efectos se le denomina “placebo”. Con métodos de instrucción, por ejemplo, ocurre que el grupo que se habrá de exponer al método innovador se ve influido por el simple hecho de decirle que se trata de un nuevo método. Lo mismo con pruebas de sabor de un producto alimenticio, programas de televisión, experiencias motivacionales, diseños arquitectónicos, sistema de calidad en las empresas, etc. Por tanto, esto debe tomarse muy en cuenta.

EJEMPLO DE EQUIVALENCIA INICIAL DE LOS GRUPOS EN EXPERIMENTOS

En el capítulo 7 del texto impreso se señaló que en un experimento los grupos de comparación deben ser equivalentes al inicio. A continuación ejemplificamos esta noción.

Supongamos que todas las variables pudieran medirse de uno a 10, la equivalencia entre grupos se conceptualizaría como en la tabla 5.1.

Tabla 5.1 Esquema para ilustrar la equivalencia inicial de los grupos

Al inicio del experimento	
Grupo 1	Grupo 2
$V_1 = 8$	$V_1 = 7.9$
$V_2 = 4$	$V_2 = 4.1$
$V_3 = 6$	$V_3 = 6$
$V_4 = 7.2$	$V_4 = 7.4$
$V_5 = 10$	$V_5 = 9.9$
20 mujeres, 21 hombres	19 mujeres, 22 hombres
Promedio de edad = 25 años	Promedio de edad = 25 años
6 meses	4 meses
$V_6 = 2$	$V_6 = 2$
$V_k = K$	$V_k = K$
$V = \text{una variable } (V_1 = \text{variable 1, } V_2 = \text{variable 2...})$	

Veamos en la tabla 5.2 un ejemplo de equivalencia entre dos grupos respecto a ciertos rasgos físicos para ilustrar el concepto.

Tabla 5.2 Ejemplo de dos grupos equivalentes

Grupo 1		Grupo 2
10 hombres de tez morena y ojos cafés		10 hombres de tez morena y ojos cafés
5 hombres de tez morena y ojos negros		4 hombres de tez morena y ojos negros
11 hombres de tez clara y ojos cafés	equivalencia	12 hombres de tez clara y ojos cafés
5 hombres de tez clara y ojos verdes	=	5 hombres de tez clara y ojos verdes
10 mujeres de pelo rubio y tez muy clara		10 mujeres de pelo rubio y tez muy clara
8 mujeres de pelo castaño claro y tez clara		9 mujeres de pelo castaño claro y tez clara
5 hombres de pelo castaño oscuro y tez clara		3 hombres de pelo castaño oscuro y tez clara

Un ejemplo de dos grupos que interculturalmente no serían equivalentes se muestra en la tabla 5.3.

Tabla 5.3 Ejemplo de dos grupos no equivalentes

Grupo 1		Grupo 2
3 venezolanos		1 venezolano
6 colombianos		3 brasileñas
5 mexicanos		2 mexicanos
6 estadounidenses		1 estadounidense
4 ingleses	≠	28 franceses
7 bolivianos		10 ingleses
3 italianos		4 rusos
5 israelitas		2 alemanes
4 afganos		5 suizos
3 cubanos		2 nicaragüenses
12 españoles		

Desde luego, es prácticamente imposible alcanzar la equivalencia perfecta o ideal, pero no deben permitirse diferencias iniciales significativas entre los grupos.

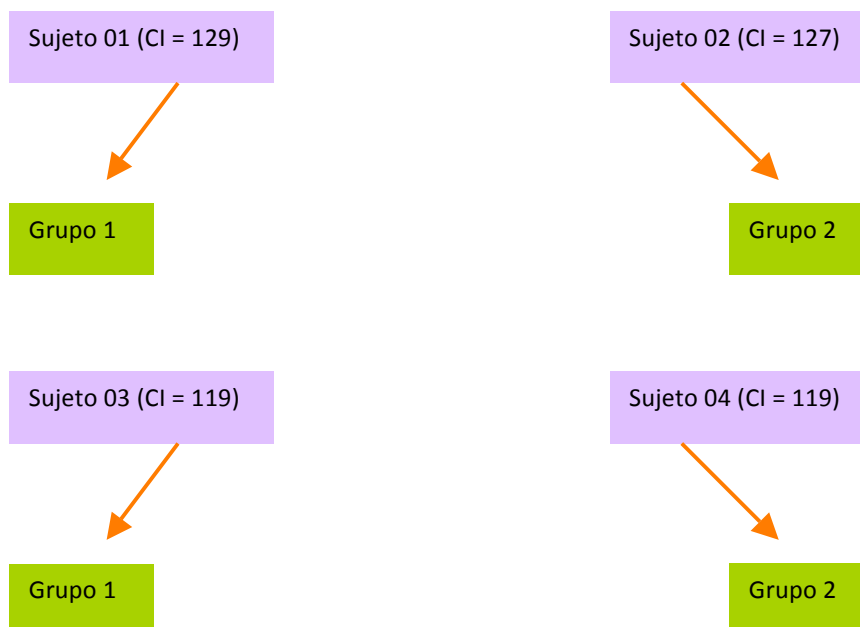
EJEMPLO DE LA TÉCNICA DE EMPAREJAMIENTO EN EXPERIMENTOS

Siguiendo lo que se expuso en el capítulo 7 del libro impreso sobre el emparejamiento, a continuación presentamos un ejemplo con dos grupos.

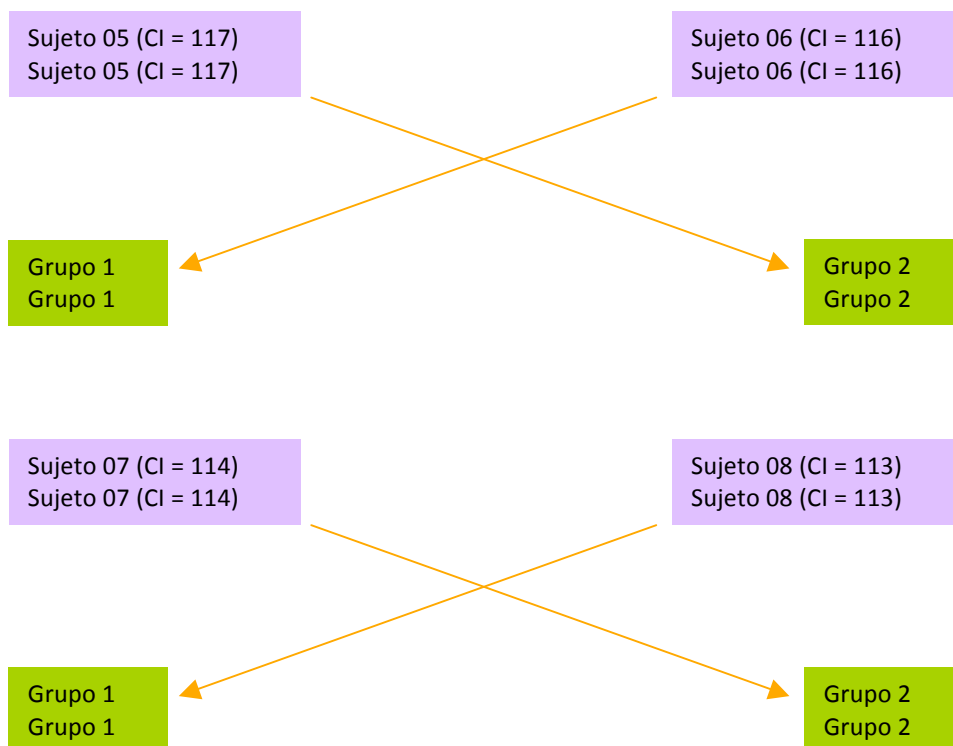
Suponga que se tuvieran 16 personas para un experimento sobre desempeño académico (recuerde la sugerencia de tener 15 o más en cada grupo; aquí se incluyen 16 únicamente para no hacer muy largo el caso); y se tuviera un grupo experimental al cual se le dará un estímulo (50% de beca para sus estudios universitarios) y un grupo de control (no se le otorga ninguna beca). La variable dependiente sería tal desempeño académico, medido por el promedio de sus calificaciones en todas las asignaturas. Los estudiantes se ordenarían de la siguiente manera:

Participante	Coeficiente de inteligencia (ci)	Participante	Coeficiente de inteligencia (ci)
01	129	09	110
02	127	10	110
03	119	11	108
04	119	12	107
05	117	13	106
06	116	14	105
07	114	15	104
08	113	16	102

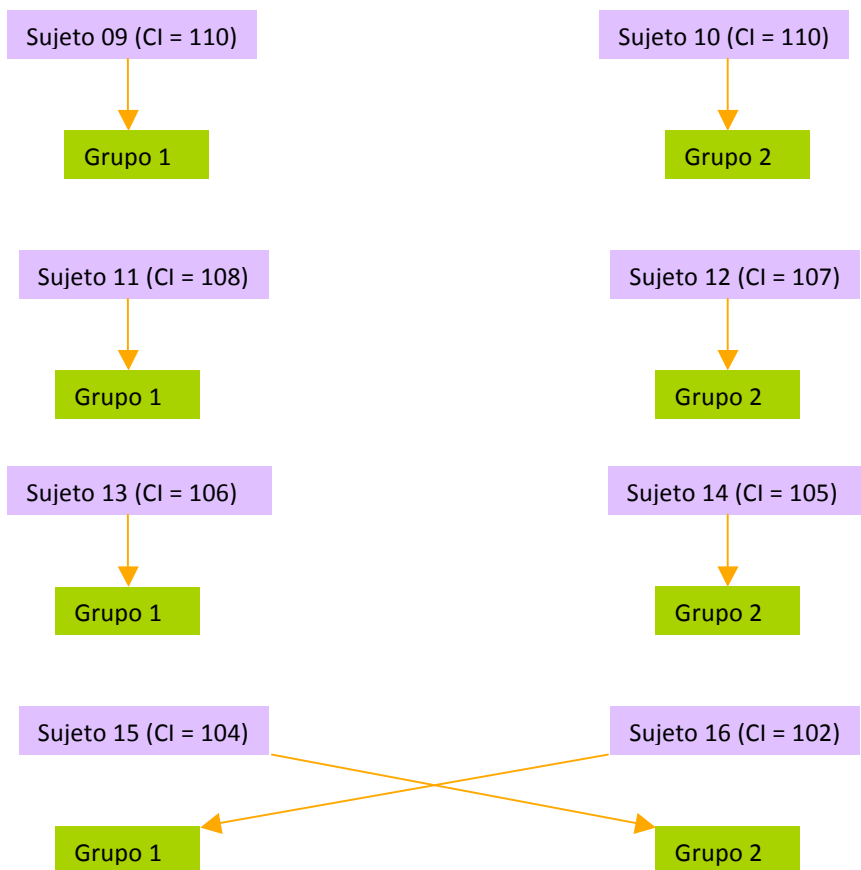
Se procede entonces a empatar gradualmente a dos grupos en la variable deseada, en este caso inteligencia:



Hasta ahora, el grupo 1 lleva dos puntos más que el grupo 2 (grupo 1 = 248, grupo 2 = 246). Hay que compensarlo.



Hasta aquí se ha conservado el balance entre los grupos, van a la par (grupo 1 = 477 y grupo 2 = 477).



Los grupos quedarían emparejados en inteligencia.

Grupo 1		Grupo 2	
<i>Sujeto</i>	<i>CI</i>	<i>Sujeto</i>	<i>CI</i>
<i>S01</i>	129	<i>S02</i>	127
<i>S03</i>	119	<i>S04</i>	119
<i>S06</i>	116	<i>S05</i>	117
<i>S08</i>	113	<i>S07</i>	114
<i>S09</i>	110	<i>S10</i>	110
<i>S11</i>	108	<i>S12</i>	107
<i>S13</i>	106	<i>S14</i>	105
<i>S16</i>	102	<i>S15</i>	104
Promedio = 112.87		Promedio = 112.87	

Son grupos equivalentes en cuanto a la variable deseada. Este procedimiento puede extenderse a más de dos grupos (si se trata de tres grupos, en lugar de parejas se va emparejando por tercias, con cuatro grupos por cuartetos, etcétera).

Respecto al emparejamiento debemos recordar lo señalado en el texto impreso: la asignación al azar es el mejor método para hacer equivalentes los grupos (más preciso y confiable). El emparejamiento no la sustituye por completo. Éste llega a suprimir o eliminar el posible efecto de la variable apareada, pero nunca nos asegura que otras variables (no apareadas) no vayan a afectar los resultados del experimento. En cambio, la asignación aleatoria garantiza que otras variables (además de la o las variables independientes de interés para el investigador) no afecten las dependientes ni confundan al experimentador. Como comentan Nunnally y Bernstein (1994), la bondad de la **asignación al azar** de los participantes o sujetos a los grupos de un diseño experimental es que el procedimiento asegura absolutamente que, en promedio, los grupos no diferirán (antes de que participen en los tratamientos experimentales) en ninguna característica más de lo que pudiera esperarse por pura casualidad.

DISEÑOS EXPERIMENTALES DE SERIES CRONOLÓGICAS MÚLTIPLES

Cuando el experimentador se interesa en analizar *efectos en el mediano o largo plazo*, porque tiene bases para suponer que la influencia de la variable independiente sobre la dependiente tarda en manifestarse. Por ejemplo, programas de difusión de innovaciones, métodos educativos o estrategias de las psicoterapias. En tales casos, *es conveniente adoptar diseños con varias pospruebas. A estos diseños*

Series cronológicas Diseños que, a través del tiempo, efectúan diversas observaciones o mediciones sobre una variable.

se les conoce como series cronológicas experimentales. En realidad, el término “serie cronológica” se aplica a cualquier diseño que efectúe varias observaciones

o mediciones sobre una variable a través del tiempo, sea o no experimental, sólo que en este caso se les llama experimentales porque reúnen los requisitos para serlo.

También *en estos diseños se tienen dos o más grupos y los participantes son asignados al azar a dichos grupos. Únicamente que, debido a que transcurre mucho más tiempo entre el inicio y la terminación del experimento, el investigador debe tener cuidado de que no ocurra algo que afecte de manera distinta a los grupos (con*

excepción de la manipulación de la variable independiente). Lo mismo sucede cuando la aplicación del estímulo lleva mucho tiempo (por ejemplo, programas motivacionales para trabajadores que duran semanas). Con el paso del tiempo es más difícil mantener la equivalencia inicial de los grupos.

Un ejemplo de cómo se diagraman las series cronológicas experimentales se muestra en la tabla 5.4.

Tabla 5.4 Ejemplos de experimentos de series cronológicas

Serie cronológica sin preprueba, con varias pospruebas y grupo de control					
RG1	X1	01	02	03	
RG2	X2	04	05	06	
RG3	X3	07	08	09	
RG4	—	010	011	012	
Serie cronológica con preprueba, con varias pospruebas y grupo de control					
RG1	01	X1	02	03	04
RG2	05	X2	06	07	08
RG3	09	—	010	011	012
Serie cronológica basada en el diseño de cuatro grupos de Solomon					
RG1	01	X	02	03	
RG2	04	—	05	06	
RG3	—	X	07	08	
RG4	—	—	09	010	

Las pospruebas pueden ser tantas como se requiera y sea posible aplicar.

Asimismo, en otras ocasiones se desea analizar la evolución de los grupos antes y después del tratamiento experimental. En esta situación pueden incluirse varias prepruebas y pospruebas, en cuyo caso se tendrían esquemas como el siguiente:

R	G_1	0	0	0	X_1	0	0	0
R	G_2	0	0	0	X_2	0	0	0
R	G_3	0	0	0	—	0	0	0

EJEMPLO

De un diseño experimental de serie cronológica

Un consultor en cuestiones organizacionales estudia el efecto sobre la dedicación en el trabajo que tiene el hecho de que se difundan una serie de valores, que la directiva de una corporación considera deben ser implantados en la empresa.

Pregunta de investigación: *¿cuanto más conozca el personal de una corporación los valores de ésta (definidos por su directiva), tendrá mayor dedicación en el trabajo?* Hipótesis de investigación: *“el mayor conocimiento de los valores corporativos genera mayor dedicación en el trabajo”.*

Hipótesis estadística: $r_{xy} > 0$

El consultor selecciona 99 personas de la corporación, de todos los niveles jerárquicos, y los asigna al azar a tres grupos:

- 1) un grupo participa en una reunión de dos horas de duración en donde se le presenta ampliamente los valores corporativos con aplicaciones a situaciones específicas de trabajo, posteriormente se le entrega un folleto con explicaciones adicionales.
- 2) Otro grupo asiste a una sesión breve donde se proporciona un folleto sobre los valores corporativos (no hay explicación verbal ni discusión o sesión de preguntas y respuestas).
- 3) El tercer grupo asiste a una sesión donde se trata algún aspecto no relacionado con el trabajo o la organización (digamos, un tema cultural de interés general).

Antes de administrar los tratamientos a todos los participantes, se les aplican tres mediciones de la dedicación en el trabajo. Asimismo, después de los tratamientos, también se les aplican tres mediciones de la misma variable (al corto, mediano y largo plazos). El diseño se diagramaría como se muestra en la figura 5.1.

Recuerde que las mediciones son de la dedicación en el trabajo. Cada grupo estaría integrado por 33 personas.

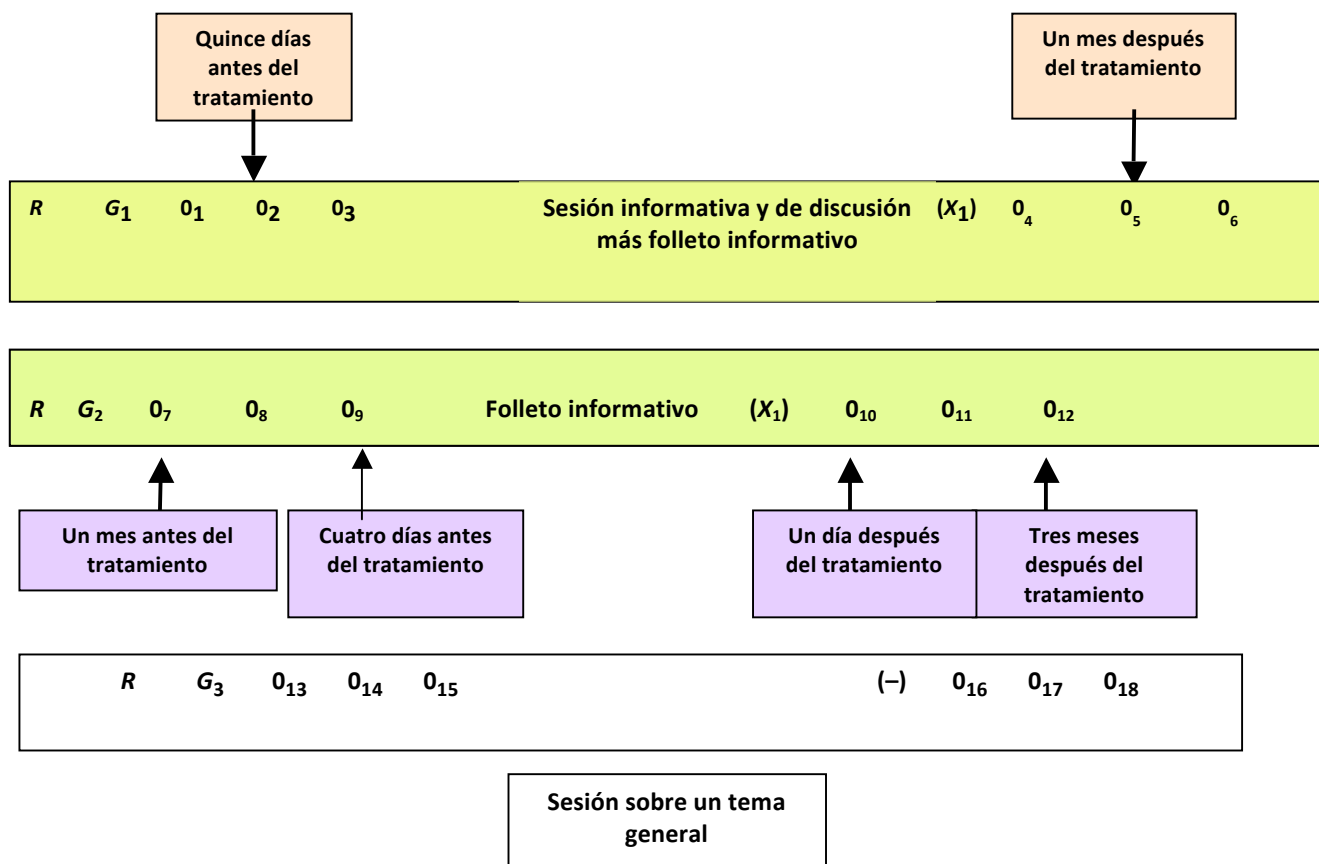
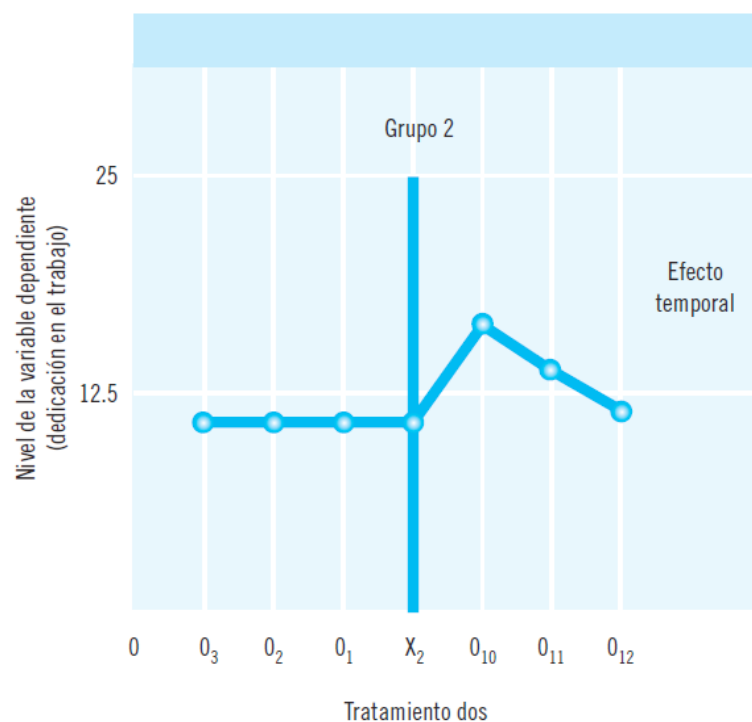
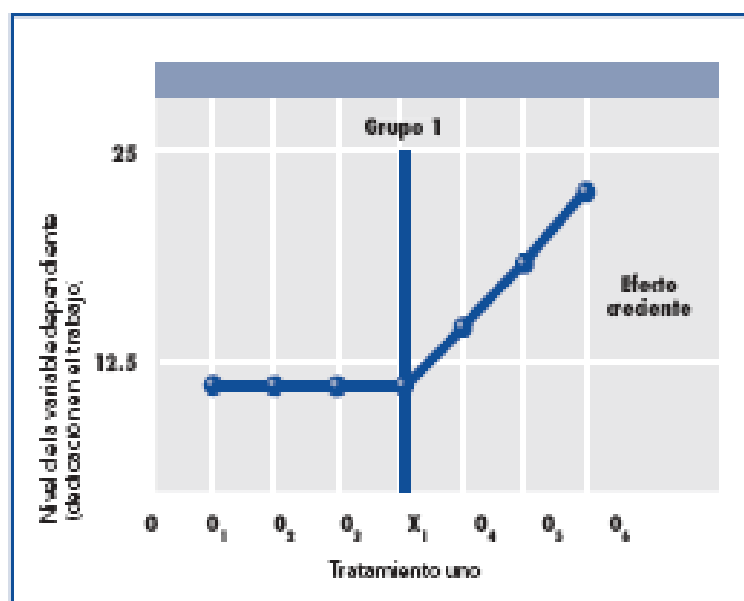


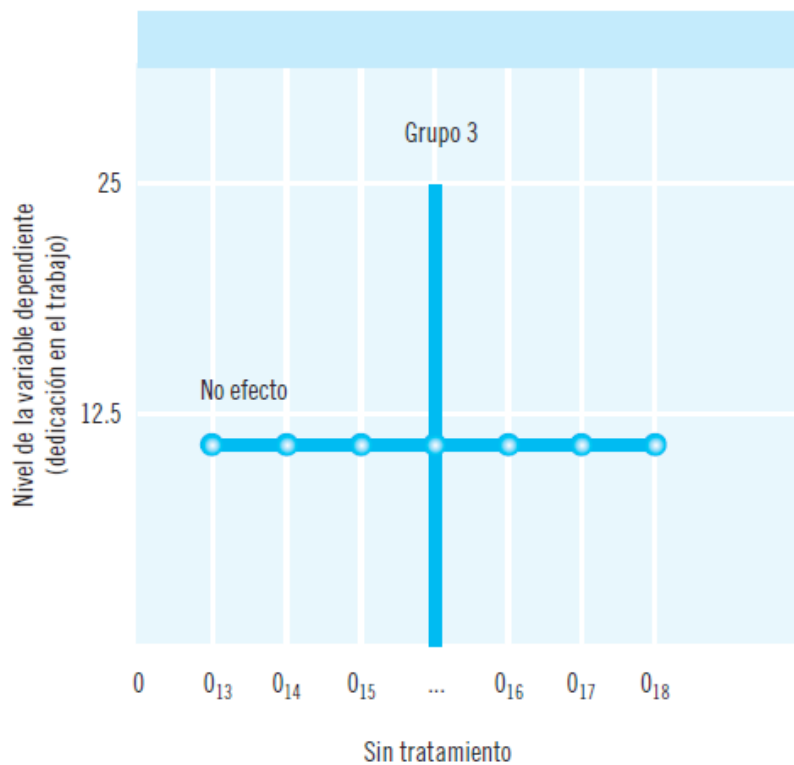
Figura 5.1 Diseño del ejemplo de serie cronológica sobre la dedicación al trabajo con tres grupos

Una ventaja del diseño consiste en que es posible evaluar la evolución comparativa de los grupos. Por ejemplo, si se encontraran los siguientes resultados con una escala (hipotética) de dedicación en el trabajo, con valores de 0 a 25:

R	G_1	11	11	11.2	X_1	16	18	21
R	G_2	10.8	11	10.9	X_2	15	14	11.8
R	G_3	11.1	10.9	11.3	—	11	10.8	11.4

Vemos que X_1 tiene un efecto que aumenta con el paso del tiempo y X_2 produce un efecto en el corto plazo, pero tiende a desvanecerse con el paso del tiempo. Esto podría graficarse del siguiente modo:





En los diseños experimentales de series cronológicas se realiza este tipo de gráficas, las cuales enriquecen la interpretación de la evolución de los grupos. Algunas de las diversas configuraciones que se presentan se muestran en las figuras 5.2, 5.3 y 5.4.

En este caso, debe observarse que no sea algún suceso ajeno el que provoque el efecto, en lugar de la manipulación de la variable independiente. Podría ocurrir que:

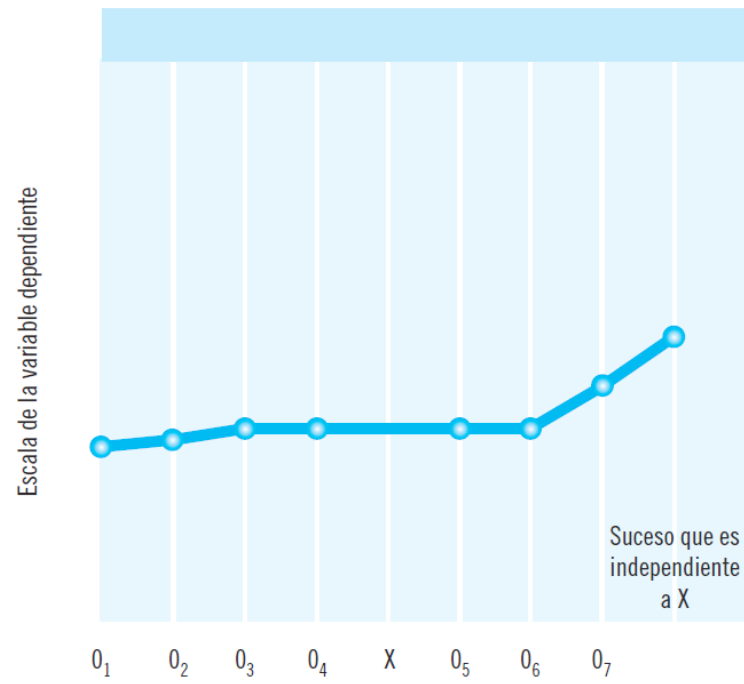


Figura 5.2 Ejemplo de un suceso ajeno a X que provoca el efecto

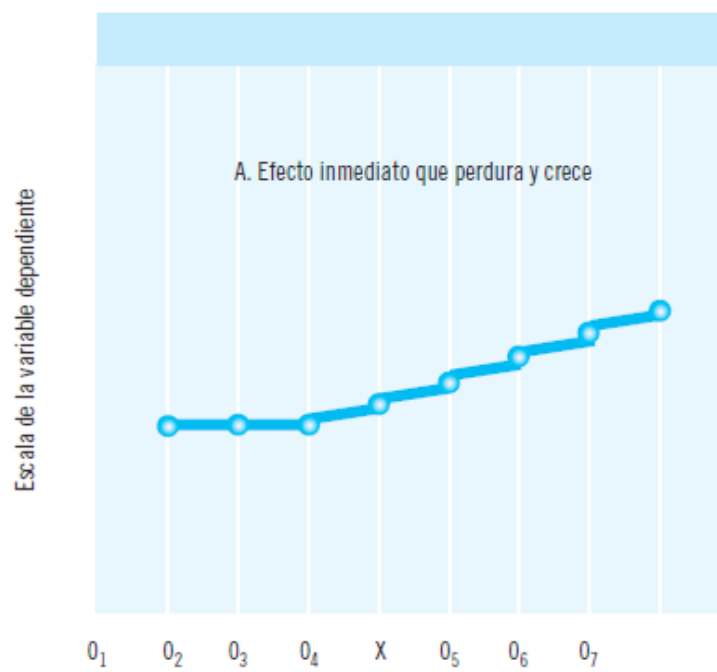


Figura 5.3 Ejemplo de un efecto que perdura

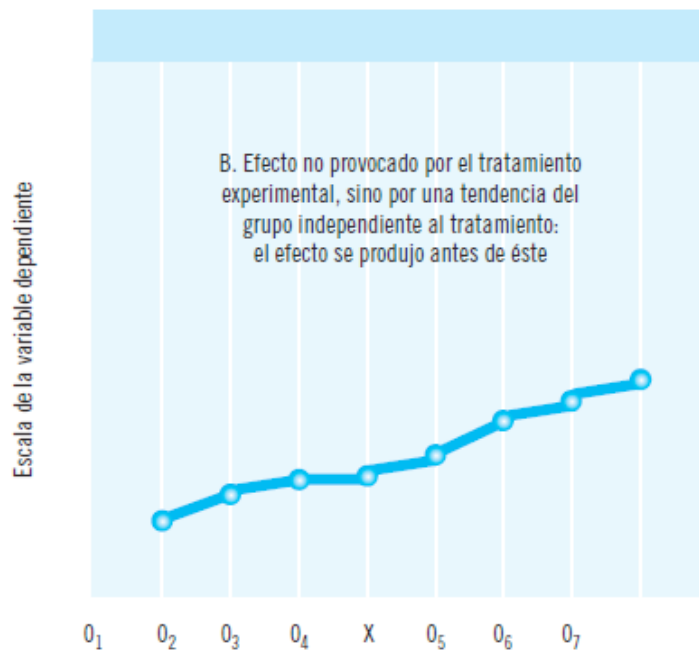


Figura 5.4 Ejemplo de ausencia de efecto por el tratamiento

Este diagrama se compararía con el del grupo de control o demás grupos para analizar lo que ocurre. Quizá se deba a que no se obtuvo una equivalencia real de los grupos al inicio del experimento.

Se tendrían tantos diagramas o gráficas como grupos, o bien podrían agruparse los efectos provocados en los distintos grupos en una sola gráfica (varios diagramas en ésta), como se observa en la figura 5.5.

Desde luego, si se mide más de una variable dependiente, en el primer caso se tendrá un diagrama para cada grupo por cada variable dependiente; mientras que en el segundo caso, un diagrama por cada variable dependiente.

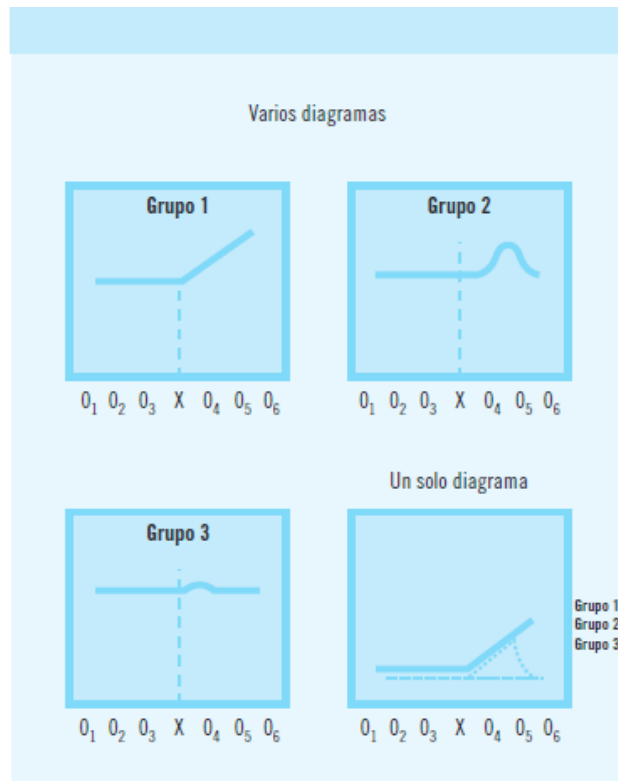


Figura 5.5 Ejemplos de gráficas de series cronológicas

EJEMPLO

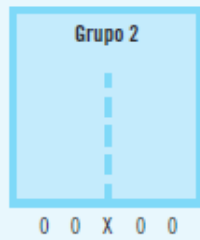
Si el tratamiento es un programa motivacional y las variables dependientes son la motivación intrínseca, la productividad y la calidad en la producción (con dos grupos experimentales y uno de control), se tendrían las opciones que se muestran en las figuras 5.6 y 5.7.

Diversos diagramas para varios grupos y variables dependientes

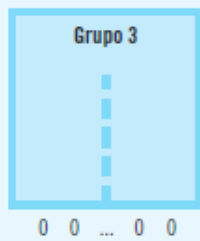
Motivación intrínseca



Motivación intrínseca



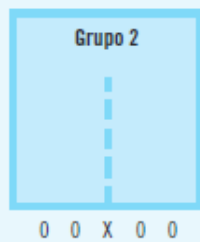
Motivación intrínseca



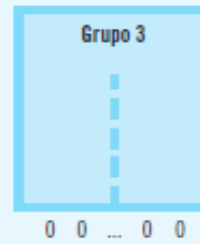
Productividad



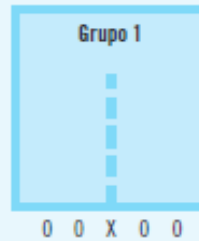
Productividad



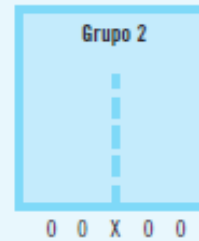
Productividad



Calidad



Calidad



Calidad

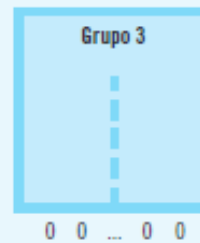


Figura 5.6 Ejemplo de varios diagramas para cada variable dependiente

Para estos diseños se suelen utilizar diversas técnicas estadísticas complejas, esto depende del nivel de medición de las variables y del tipo de análisis e interpretación que se desee realizar, como *análisis de regresión múltiple* o *análisis de cambio y series de tiempos*. Se recomiendan las siguientes fuentes para conocer tales análisis: Sugerimos a Cryer y Chan (2009); Hawkins (2008); Box, Jenkins y Reinsel (2008); Kollo y von Rosen (2005), y Kerlinger y Pedhazur (1997).

Diseños de series cronológicas con repetición del estímulo

En ocasiones, el investigador anticipa que el tratamiento o estímulo experimental *no* tiene efecto o es mínimo si se aplica una sola vez, tal como sería hacer ejercicio físico un sólo día (no se esperaría un cambio en la musculatura); o como sería consumir vitaminas por una única vez. También en ocasiones el investigador quiere conocer el efecto sobre las variables dependientes cada ocasión que se aplica el estímulo experimental.

Por ejemplo, en técnicas de condicionamiento es común que uno se cuestione: ¿cuántas veces debo aplicar el reforzamiento a una conducta para lograr condicionar la respuesta a un estímulo? *En estos casos es posible repetir el tratamiento experimental y administrar una posprueba después de cada aplicación*, para evaluar su efecto.

Los participantes se asignan al azar a los distintos grupos, a cada grupo se le administra varias veces el tratamiento experimental que le corresponde. Algunos de estos diseños diagramados se muestran en la tabla 5. 5.

Tabla 5.5 Ejemplos de diseños cronológicos con repetición del estímulo

El mismo tratamiento se aplica dos veces al grupo experimental:										
R	G ₁	0 ₁	X ₁	0 ₂	X ₁	0 ₃				
R	G ₂	0 ₄	—	0 ₅	—					0 ₆
Cada tratamiento se aplica cuatro veces al grupo respectivo:										
R	G ₁	0 ₁	X ₁	0 ₂	X ₁	0 ₃	X ₁	0 ₄	X ₁	0 ₅
R	G ₂	0 ₆	X ₂	0 ₇	X ₂	0 ₈	X ₂	0 ₉	X ₂	0 ₁₀

R	G ₃	0 ₁₁	X ₃	0 ₁₂	X ₃	0 ₁₃	X ₃	0 ₁₄	X ₃	0 ₁₅		
R	G ₄	0 ₁₆	—	0 ₁₇	—	0 ₁₈	—	0 ₁₉	—	0 ₂₀		
<i>En algunos casos se prescindiría de las pruebas, y el experimentador, por alguna justificación teórica o empírica, podría aplicar pospruebas a intervalos sistemáticos diferentes.</i>												
R	G ₁	X ₁	0 ₁	X ₁	X ₁	X ₁	0 ₂	X ₁	X ₁	X ₁	0 ₃	
R	G ₂	X ₂	0 ₄	X ₂	X ₂	X ₂	0 ₅	X ₂	X ₂	X ₂	0 ₆	
R	G ₃	—	0 ₇	—	—	—	0 ₈	—	—	—	0 ₉	
<i>O bien, aplicar las pospruebas a intervalos irregulares (por alguna determinada razón):</i>												
R	G ₁	X ₁	X ₁	0 ₁	X ₁	0 ₂	X ₁	X ₁	X ₁	0 ₃	X ₁	0 ₄
R	G ₂	X ₂	X ₂	0 ₅	X ₂	0 ₆	X ₂	X ₂	X ₂	0 ₇	X ₂	0 ₈
R	G ₃	—	—	0 ₉	—	0 ₁₀	—	—	—	0 ₁₁	—	0 ₁₂

Un ejemplo de estos diseños sería el caso de un publicista que pretende analizar los efectos de un comercial televisivo en la preferencia del producto anunciado, en comparación con otras marcas, y postula la hipótesis de que una sola exposición al comercial no surtirá efecto alguno.

Las pruebas estadísticas usuales para estos diseños son las mismas que para las series cronológicas múltiples.

Diseños con tratamientos múltiples

A veces el investigador desea analizar *el efecto de la aplicación de los distintos tratamientos experimentales a todos los grupos o participantes*. En estos casos es posible utilizar los diseños con tratamientos múltiples. La aplicación de tratamientos puede ser individual o en un grupo y hay distintas variaciones:

a) *Varios grupos*. En este caso, se tienen varios grupos a los cuales se asignan los sujetos o participantes al azar. A cada grupo se le aplican todos los tratamientos. La secuencia de la aplicación de tratamientos puede o no ser la misma para todos los grupos y es posible administrar una o más pospruebas a los grupos (posteriores a cada

tratamiento experimental). Dos diagramas de estos diseños son los de la tabla 5. 6.

Misma secuencia para los grupos									Secuencia diferente								
R	G ₁	X ₁	0 ₁	X ₂	0 ₂	X ₃	0 ₃		R	G ₁	X ₁	0 ₁	X ₂	0 ₂	X ₃	0 ₃	
R	G ₂	X ₁	0 ₄	X ₂	0 ₅	X ₃	0 ₆		R	G ₂	X ₂	0 ₄	X ₃	0 ₅	X ₁	0 ₆	
R	G ₃	X ₁	0 ₇	X ₂	0 ₈	X ₃	0 ₉		R	G ₃	X ₃	0 ₇	X ₂	0 ₈	X ₁	0 ₉	

Tabla 5.6 Diseños con tratamientos múltiples, varios grupos

Con secuencia diferente. El experimentador debe tener cuidado al interpretar las segundas pospruebas y mediciones subsecuentes, ya que quizás exista una influencia diferente en los grupos provocada por distintas secuencias de los tratamientos. De hecho, durante el experimento es muy probable que haya diferencias entre grupos, incluso al finalizar el experimento los resultados se deban, en buena medida, a la secuencia con que fueron administrados los tratamientos.

Los diseños experimentales con tratamientos múltiples y secuencia diferente en los grupos, así como los dos casos que vamos a ver a continuación, llegan a tener distintos efectos que deben analizarse con minuciosidad. Algunos tratamientos tienen efectos reversibles; en esta situación no hay interferencia entre tratamientos y las pospruebas se ven influidas únicamente por el tratamiento inmediato anterior (por ejemplo 0₃, del diseño con secuencia diferente, se vería afectada por X₃, pero no por X₂ o X₁), ello facilita la interpretación. Pero a menudo los efectos no son reversibles, sino aditivos o interactivos; esto es, los resultados de una posprueba se pueden ver influidos no sólo por el tratamiento inmediatamente anterior, sino por los que antecedieron a éste, y no es fácil saber cuánto se debió a X₁, cuánto a X₂, o X_k. Para ello, en el análisis debe incluirse la secuencia como factor.

b) *Un solo grupo.* En situaciones donde sólo se cuenta con un número reducido de participantes para el experimento, es posible realizar un diseño con tratamientos

múltiples y un solo grupo. No hay asignación al azar puesto que se tiene a un único grupo. La equivalencia se obtiene puesto que no hay nada más similar a un grupo que este mismo. El grupo hace las veces de “grupos experimentales” y de “control”. Este diseño se diagrama así:

$G \text{ único } X_1 \ 0_1 \ X_2 \ 0_2 \text{ — } 0_3 \ X_3 \ 0_4 \text{ — } 0_5 \ X_k \ 0_k \dots$

Cuando se considere conveniente, se utiliza como grupo de control, por ejemplo, antes de 0_3 y 0_5 . Sin embargo, tal diseño está limitado a que los efectos de los tratamientos múltiples sean reversibles; de lo contrario no es un diseño experimental, sino cuasiexperimental. Si en estos diseños se introduce sistemáticamente y como variable independiente la secuencia de administración de los tratamientos, se convierten en factoriales (que se verán a continuación).

Las pruebas estadísticas que se utilizan en estos diseños son las mismas que se tienen para las series cronológicas y los diseños con repetición del estímulo.

Hasta aquí se han revisado diseños experimentales que manipulan una sola variable independiente (los tratamientos representan niveles de presencia o manipulación de ésta; X_1 , X_2 , X_3 , X_k son variaciones de la misma variable independiente; al menos como se han concebido en este libro), de este modo, miden una o más variables dependientes. Por ejemplo, en un experimento con métodos educativos, en lugar de medir nada más el aprendizaje, es posible medir además la motivación del alumno, su integración al grupo, etc.; esto sería en la preprueba y la posprueba. Normalmente, se miden diversas variables dependientes para optimizar el costo, al analizar los efectos de la variable independiente sobre cada dependiente. En ocasiones, como parte de las pospruebas, se incluyen mediciones para verificar qué tanto funcionó la manipulación (verificaciones de la manipulación) o a veces estas verificaciones son independientes de las pospruebas.

Con el fin de ver si las secuencias se relacionan con características de los grupos, algunos autores consideran que introducir sistemáticamente otros elementos a los diseños, tales como presencia-ausencia de la preprueba (por ejemplo, el diseño de cuatro grupos de Solomon, visto en el capítulo 7 del texto impreso), así como secuencias diferentes de tratamientos en varios grupos. Por ejemplo, tener un diseño así:

Niños de 8º año,	G1	X1	01	X2	02
	G2	X2	03	X1	04
Niños de 6º año,	G3	X1	05	X2	06
	G4	X2	07	X1	08
Niños de 4º año,	G5	X1	09	X2	010
	G6	X2	011	X1	012

O un número de pospruebas diferentes en los grupos. Por ejemplo, con un diseño como el siguiente, se tienen diseños factoriales:

R	G1	X1	01	02
R	G2	X1	01	
R	G3	X2	01	02
R	G4	X2	01	
R	G5	—	01	02
R	G6	—	01	

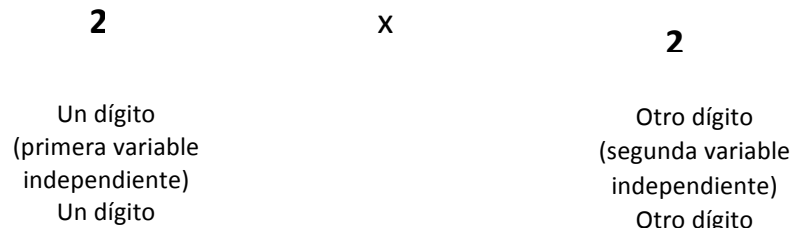
De hecho, estos autores tienen razón, porque se manipulan como si fueran una variable independiente. Sin embargo, la experiencia ha demostrado a algunos profesores que los alumnos que ven por primera vez el tema suelen desconcertarse si se analizan como diseños factoriales, a menos que éstos se expliquen primero.

DISEÑOS FACTORIALES

Los diseños factoriales manipulan dos o más variables independientes e incluyen dos o más niveles de presencia en cada una de las variables independientes. Se utilizan muy a menudo en la investigación del comportamiento. La construcción básica de un diseño factorial consiste en que todos los niveles de cada variable independiente sean tomados en combinación con todos los niveles de las otras variables independientes (Wiersma y Jurs, 2008).

Diseño factorial 2 X 2

El diseño factorial más simple manipula (hace variar) dos variables, cada una con dos niveles. A este diseño se le conoce como “diseño factorial 2 X 2”, en donde el número de dígitos indica el número de variables independientes:



El valor numérico de cada dígito indica el número de niveles o modalidades de la variable independiente en cuestión. En este caso es “2”, lo cual quiere decir que cada una de las variables tiene dos niveles. Como mencionan Wiersma y Jurs (2008), no es necesario que los valores numéricos sean los mismos para todas las variables independientes. En teoría, puede haber cualquier número de variables independientes con cualquier número de niveles cada una. Por ejemplo, el diseño factorial 2 X 2 X 3 indica que hay tres variables independientes, la primera y la segunda con dos niveles, mientras que la tercera con tres niveles. El diseño factorial 4 X 5 X 2 X 3 indica una variable independiente con cuatro niveles, otra con cinco, otra más con dos y una última con tres.

Un ejemplo de un diseño factorial 2 X 2 sería tener como variables independientes “método de enseñanza” y “género”. La primera con dos niveles: “método de enseñanza tradicional-oral” y “método de enseñanza por medio de video”. La segunda con los niveles “masculino” y “femenino”.

Otros diseños factoriales

El número de grupos que se forman en un diseño factorial es igual a todas las posibles combinaciones que surjan al cruzar los niveles de una variable independiente con los niveles de las otras variables independientes. Así, en un diseño 2 X 2 tendremos cuatro grupos ($2 \text{ por } 2 = 4$); en un diseño 3 X 2 resultarán seis grupos, y en un diseño 3 X 3 X 3, 27 grupos. Debe observarse que el resultado de

la multiplicación es el número de grupos resultante. En estos diseños, el número de grupos aumenta con rapidez con el incremento del número de variables independientes o niveles (exponencialmente). Veámoslo:

2×2	=	4	}	Número de grupos resultantes o del experimento.
2×3	=	6		
3×3	=	9		
3×4	=	12		
$3 \times 2 \times 2$	=	12		
$3 \times 3 \times 4$	=	36		

Ello se debe a que los niveles tienen que tomarse en cuenta respecto a todas sus posibles combinaciones.

Wiersma y Jurs (2008) señalan que en los diseños experimentales factoriales, *al menos una de las variables independientes debe ser experimental*; las demás pueden ser variables orgánicas, introducidas en el diseño con fines de control (por ejemplo, género, edad, año, escolaridad, inteligencia en intervalos, etcétera).

Diseño 2 X 2

		Variable independiente A	
		A_1	A_2
Variable independiente B	B_1	A_1B_1	A_2B_1
	B_2	A_1B_2	A_2B_2

Figura 5.8 Un diseño factorial 2 X 2

Con objeto de simplificar la forma en que se diagraman los diseños factoriales,

acudiremos a la simbología que comúnmente se utiliza.¹ Para designar a las variables independientes se usan letras ($A, B, C, \dots K$) y para los niveles, números (1, 2, 3, $\dots K$); las combinaciones de letras y números que aparecen en las casillas (o celdas) representan las mezclas de niveles de las variables independientes. Cada celda es un grupo. En la figura 5.8 (en la página anterior) se diagrama un diseño factorial 2 X 2. Otro ejemplo sería un diseño factorial 2 X 4 X 3 (figura 5.9).

Diseño 2 X 4 X 3

	A					
	A₁			A₂		
	C₁	C₂	C₃	C₁	C₂	C₃
B₁	A ₁ B ₁ C ₁	A ₁ B ₁ C ₂	A ₁ B ₁ C ₃	A ₂ B ₁ C ₁	A ₂ B ₁ C ₂	A ₂ B ₁ C ₃
B₂	A ₁ B ₂ C ₁	A ₁ B ₂ C ₂	A ₁ B ₂ C ₃	A ₂ B ₂ C ₁	A ₂ B ₂ C ₂	A ₂ B ₂ C ₃
B₃	A ₁ B ₃ C ₁	A ₁ B ₃ C ₂	A ₁ B ₃ C ₃	A ₂ B ₃ C ₁	A ₂ B ₃ C ₂	A ₂ B ₃ C ₃
B₄	A ₁ B ₄ C ₁	A ₁ B ₄ C ₂	A ₁ B ₄ C ₃	A ₂ B ₄ C ₁	A ₂ B ₄ C ₂	A ₂ B ₄ C ₃

Figura 5.9 Diseño 2 X 4 X 3

Observe en la figura 5.9 que todas las posibles combinaciones de los niveles entre A , B y C están presentes; además, ninguna combinación es exactamente igual a la otra. Cada combinación representa una celda o un grupo. Si las tres variables se manipulan deliberadamente (recordemos que al menos una debe serlo para que hablemos de experimento), los participantes tienen que asignarse al azar a todas las celdas o grupos. Si dos variables se manipulan intencionalmente (por ejemplo, B y C), los sujetos de cada nivel de la variable restante (A) serán asignados al azar a las casillas que les corresponde. Veámoslo con un ejemplo. Si A = género (A_1 , masculino;

¹ Por ejemplo, Christensen (2006); y Wiersma y Jurs (2008).

A_2 , femenino), B = violencia televisada (B_4 , elevada; B_3 , mediana; B_2 , baja, y B_1 , nula) y C = orientación paternal sobre el programa visto (C_1 , por parte de ambos padres; C_2 , del padre, y C_3 , de la madre). Si hay 120 niños y 120 niñas, los niños (A_1) se asignarían al azar a las celdas donde A_1 está presente (10 niños en cada celda), y las niñas (A_2) a las 12 casillas restantes (donde A_2 está presente, 10 niñas por celda). Si una sola variable es la que se manipula deliberadamente (C por ejemplo), los participantes de los niveles combinados de las otras dos variables se asignan al azar a los niveles de aquélla (C_1 , C_2 y C_3 , en el ejemplo). Los sujetos $A_1 B_1$ serían asignados aleatoriamente a C_1 , C_2 y C_3 , igual los sujetos $A_1 B_2$, $A_1 B_3$, etcétera.

En los diseños factoriales es posible agregar un grupo de control o varios (que no se expongan a la variable o las variables manipuladas deliberadamente, figura 5.10).

Diseño 2 X 2

	A_1	A_2	
B_1	$A_1 B_1$	$A_2 B_1$	
B_2	$A_1 B_2$	$A_2 B_2$	
			control

Figura 5.10 Diseño factorial con grupo de control

Utilidad de los diseños factoriales

Los diseños factoriales son sumamente útiles porque *permiten al investigador evaluar los efectos de cada variable independiente sobre la dependiente por separado, así como los efectos de las variables independientes de manera conjunta*. Por medio de estos diseños se observan los *efectos de interacción entre las variables independientes*.

En términos de Wiersma y Jurs (2008), la interacción es un efecto producido sobre la variable dependiente, de tal manera que el efecto de una variable

independiente deja de permanecer constante en los niveles de la otra u otras independientes. El efecto de interacción está presente si el efecto conjunto de las variables independientes no es igual a sus efectos por separado (aditivos). Ello significa que el efecto de una variable independiente por sí mismo no es igual que cuando se toma en combinación con los niveles de otra variable independiente. Por ejemplo, si el alto contenido de violencia televisada afecta sólo cuando hay orientación sobre el programa por parte de la madre, pero no cuando dicha orientación está a cargo del padre o de ambos.

Así, *hay dos tipos de efectos* que es posible evaluar en los diseños factoriales: *los efectos de cada variable independiente* (llamados efectos principales) y *los efectos de interacción entre dos o más variables independientes* (si se tienen cuatro variables, por ejemplo, pueden interactuar dos entre sí y otras dos entre sí, o pueden interactuar tres o las cuatro variables independientes). Los diseños factoriales responden a estructuras entre variables que se esquematizan en la siguiente forma:

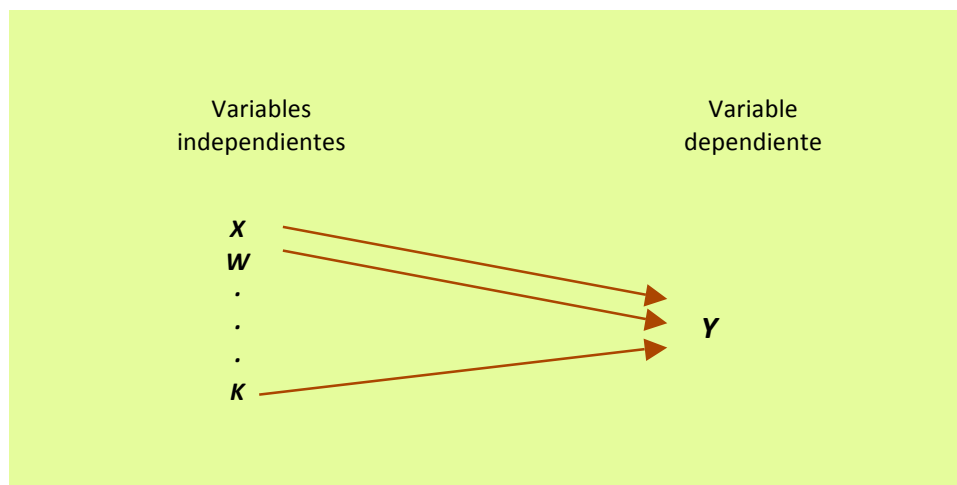
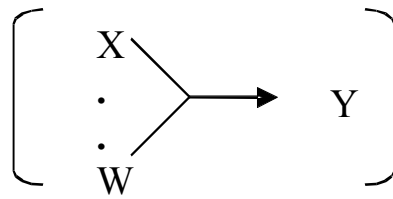
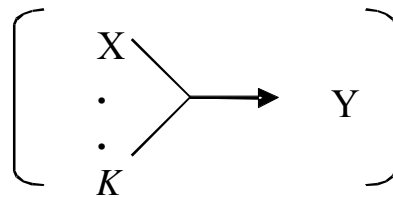


Figura 5.11 Ejemplo de estructura de un diseño factorial

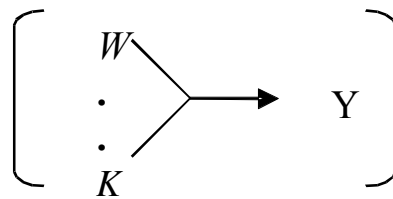
Asimismo, se analizan el efecto de X sobre Y ($X \rightarrow Y$), de W sobre Y ($W \rightarrow Y$) y el de ... K sobre Y ($K \rightarrow Y$); el efecto conjunto de X y W sobre Y .



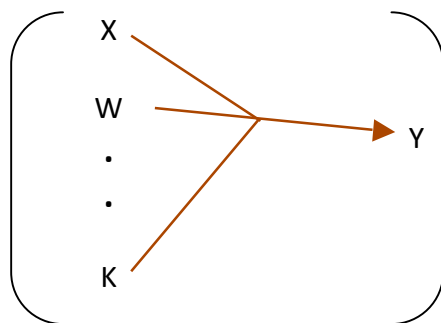
El efecto conjunto de X y W sobre Y .



El efecto conjunto de W y K sobre Y .



Finalmente el efecto conjunto de todas las variables independientes sobre Y (tres variables).



EJEMPLO

Supongamos una investigación que tiene como hipótesis: “a mayor exposición por

parte de los adolescentes a videos musicales con alto contenido sexual, habrá una mayor predisposición para establecer contacto heterosexual”.

Posteriormente se diseña un experimento para someterla a prueba. La variable independiente es la *exposición a contenidos sexuales* (por medio de la televisión) y la dependiente es la *predisposición para establecer contacto sexual*. Se decide agregar otra variable independiente: género. Entonces se tiene un diseño factorial con dos variables independientes. La exposición tendría tres niveles: *a)* contenido sexual elevado y manifiesto, *b)* contenido sexual moderado y *c)* contenido “romántico”.

Cada video que se elabore tendrá, entonces, tres versiones: la música es la misma, los modelos, la duración, la historia y el contexto también. La única diferencia es el tratamiento sexual de los contenidos verbales y no verbales. El género implicaría dos niveles: masculino y femenino.

El esquema del diseño se indica en la figura 5.12.

Exposición al contenido sexual

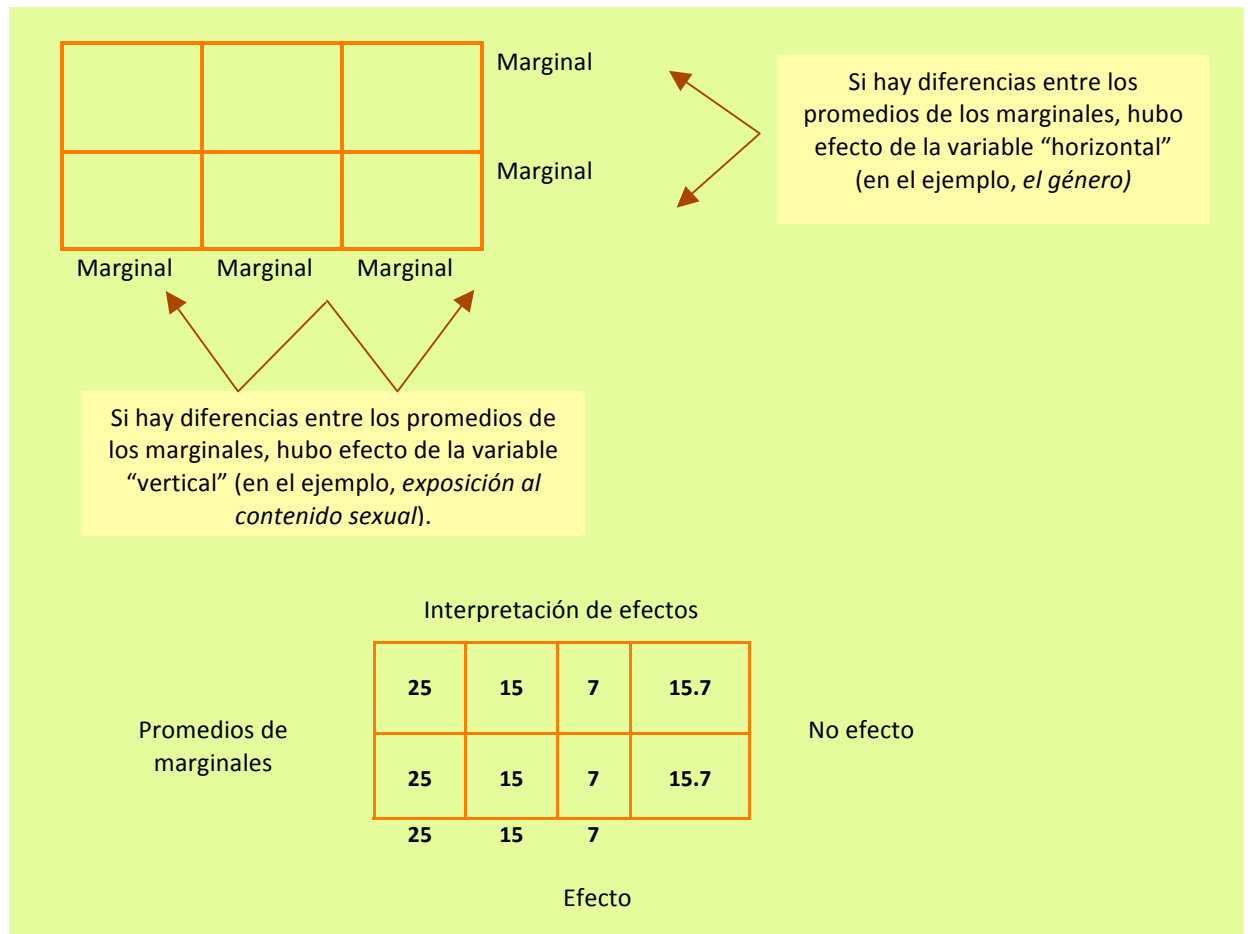
Género	Exposición al contenido sexual		
	Elevada	Moderada	Mínima
Masculino			
Femenino			

Figura 5.12 Diseño factorial 2 X 3 (correspondiente al ejemplo)

Christensen (2006) desarrolla una excelente explicación de los efectos principales e interactivos en los diseños factoriales, la cual sirve de base para la exposición de dichos efectos que se incluye a continuación.

Supongamos que, una vez que los grupos se han expuesto a los videos, se les aplica una posprueba que mide su predisposición para establecer contacto sexual y se obtiene el promedio de cada grupo (asimismo, pensemos que los resultados que lleguen a obtenerse en esta prueba oscilaran entre 0 y 30, donde un valor mayor indica una más alta predisposición). Analizamos varias configuraciones posibles de

resultados (figura 5.13).



A

	Elevada	Moderada	Mínima
Masculino	25	15	7
Femenino	25	15	7

La variable exposición tiene efecto; la variable género no tiene efecto. No hay efecto de interacción.

B

	Elevada	Moderada	Mínima
Masculino	18	18	18
Femenino	8	8	8

La variable exposición no tiene efecto; la variable género tiene efecto. No hay efecto de interacción.

C

	Elevada	Moderada	Mínima
Masculino	26	18	10
Femenino	10	18	25.9

Ninguna de las variables tiene efecto por sí misma, pero hay un efecto de interacción.

D

	Elevada	Moderada	Mínima
Masculino	23	18	13
Femenino	18	13	8

Las dos variables tienen efecto, pero no hay efecto de interacción

E

	Elevada	Moderada	Mínima
Masculino	29	23	12
Femenino	20	12	12

Las dos variables tienen efecto y hay efecto de interacción.

F

	Elevada	Moderada	Mínima
Masculino	8	7.9	8.1
Femenino	7.9	8	8

No hay ningún tipo de efecto.

Figura 5.13 Diagramas del ejemplo de diseño factorial

Interpretación y posibles configuraciones

En la configuración *A*, *no hay diferencias por género* (cada casilla en cada nivel de género es igual a la del otro nivel). En cambio, *hay diferencias entre los niveles de exposición en ambos géneros*. Los *efectos principales* (es decir, los de cada variable por separado) se observan comparando los promedios de los marginales de los niveles de cada variable (vea la interpretación de efectos de la figura 5.13).

En la configuración *B*, *no hay diferencias por exposición, pero sí por género*.

En la configuración *C*, las diferencias entre las celdas se explican porque *las dos variables interactúan* (cuando el género es masculino y la exposición elevada, asimismo cuando el género es femenino y la exposición mínima, se obtiene un valor; cuando el género es masculino y la exposición mínima, y cuando el género es femenino y la exposición elevada, se obtiene otro valor; finalmente, cuando ambos géneros se exponen moderadamente se obtiene un valor distinto de las demás celdas). El máximo efecto se encuentra en “masculino” y “elevada” y “femenino” y “mínima” (26 y 25.9). *No hay efectos principales*.

En la configuración *D*, *hay cambios verticales y horizontales*, provocados por *efectos principales, pero no efecto de interacción* (ambas variables tienen efecto por sí mismas, únicamente).

En la configuración *E*, hay *efectos principales* (cada variable por sí misma afecta) y también *efecto de interacción* (éste se alcanza si la diferencia entre las medias de los niveles de variación de una variable independiente cambian en función de los niveles de variación de la otra variable independiente, como también ocurrió en la configuración *C*).

En la configuración *F*, las diferencias entre todas las celdas es prácticamente nula: no hay ninguna clase de efecto.

Métodos estadísticos de los diseños factoriales

Los métodos estadísticos más usuales para estos diseños son *el análisis de varianza factorial* (ANOVA) y *el análisis de covarianza* (ANCOVA), con la variable dependiente medida en intervalos, y la *chi-cuadrada* (c^2) para múltiples grupos, con esa variable medida nominalmente.

Por último, a estos diseños se les pueden agregar más variables dependientes (tener dos o más) y se convierten en diseños multivariados experimentales que utilizan como método estadístico el *análisis multivariado de varianza* (MANOVA).

¿QUÉ OTROS EXPERIMENTOS EXISTEN?: CUASIEXPERIMENTOS

Ya se comentó, en el capítulo 7 del libro, que los diseños cuasiexperimentales también manipulan deliberadamente una o más variables independientes para observar su efecto y relación con una o varias dependientes, sólo que trabajan con “grupos intactos”, formados por motivos ajenos al experimento: *en los diseños cuasiexperimentales los participantes no se asignan al azar a los grupos ni se emparejan, sino que dichos grupos ya estaban integrados previamente al experimento*. Por ejemplo, si los grupos del experimento son cuatro grupos de diferentes áreas de una empresa que estaban conformados con anterioridad al experimento (las razones por las que cada participante se integró a cada área no tienen que ver con la realización del experimento). Veámoslo gráficamente:

Grupo A	(50 empleados del departamento de <i>producción</i>)	►	Grupo experimental con X_1
Grupo B	(26 empleados de la gerencia de <i>recursos humanos</i>)	►	Grupo experimental con X_2
Grupo C	(37 empleados de la gerencia de <i>administración y finanzas</i>)	►	Grupo experimental X_3
Grupo D	(29 empleados del departamento de <i>aseguramiento de la calidad</i>)	►	Grupo de control

Problema de los diseños cuasiexperimentales

Estos diseños *se utilizan cuando no es posible asignar en forma aleatoria los participantes a los grupos que recibirán los tratamientos experimentales*. La falta de aleatorización introduce posibles *problemas de validez interna y externa*. Como comenta Weiss (1990, p. 89):

[...]estos diseños deben luchar con la selección como fuente posible de interpretación equivocada, lo mismo que con la interacción de la selección y otros factores; así como, posiblemente, con los efectos de la regresión.

Asimismo, diversos elementos pudieron operar en la formación de los grupos (que no están bajo el control del investigador), que impiden afirmar que éstos son representativos de poblaciones más amplias. De este modo, dado que su validez es menor que la de los experimentos “puros”, reciben el nombre de cuasiexperimentos.

A causa de los problemas potenciales de validez interna, en estos diseños el investigador debe intentar establecer la semejanza entre los grupos; esto requiere considerar las características o variables que estén relacionadas con las variables estudiadas (Wiersma y Jurs, 2008; Babbie, 2009). Por ejemplo, si grupos intactos de trabajadores se involucran en un experimento sobre motivación, el turno probablemente tenga que ser introducido como una constante (grupos intactos, que sean todos del mismo turno) o como otra variable independiente (de control). Asimismo, el investigador deberá buscar evidencia de que los grupos son equiparables en salario, productividad, competencia, antigüedad en la organización y, en general, en todo lo que genere diferencias entre los grupos. Cuanta mayor información se obtenga sobre los grupos, mayores bases se tendrán para establecer su semejanza. En algunos casos se observará si hay la misma proporción de mujeres y hombres en los grupos, si la edad promedio es similar, si los grupos no fueron constituidos con base en un criterio que pudiera afectar (por ejemplo, formación de los salones por inteligencia) y si a los grupos en el pasado no les ha ocurrido algo que llegara a influir en los resultados. Además, como mencionan Campbell y Stanley (1966, p. 70):

...precisamente porque hay falta de control experimental total, es imprescindible que el investigador conozca a fondo cuáles son las variables particulares que su diseño específico no controla. Así, estará más pendiente de su posible influencia y tendrá mejores elementos para evaluarla.

La ausencia de asignación al azar hace que se ponga especial atención al interpretar los resultados y se tenga sumo cuidado de no caer en interpretaciones erróneas. Las limitaciones deben identificarse con claridad, la equivalencia de los grupos debe discutirse y la posibilidad de generalizar los resultados, así como la representatividad, deberán argumentarse sobre una base lógica (Wiersma y Jurs, 2008).

Los cuasiexperimentos difieren de los experimentos “puros” en la equivalencia inicial de los grupos (los primeros trabajan con grupos intactos y los segundos utilizan un método para hacer equivalentes a los grupos). Sin embargo, esto *no*

quiere decir que sea imposible tener un caso de cuasiexperimento, donde los grupos sean equiparables en las variables relevantes para el estudio. Si así fuera, los cuasiexperimentos ya se hubieran desechado como diseños de investigación. Más bien quiere decir que, en algunos casos, los grupos pueden diferir significativamente o no ser equiparables; de este modo, el investigador debe analizar si los grupos son

Cuasiexperimento Experimento en el que los sujetos no se asignan al azar a los grupos ni se emparejan, porque tales grupos ya existían. La variable exposición no tiene efecto;

o no comparables desde el inicio. En esta última situación el investigador debe declinar la realización de la investigación con fines explicativos y limitarse a propósitos descriptivos y/o correlacionales.

Una recomendación con el fin de consolidar la validez interna de esta clase de diseños experimentales, es que los tratamientos o estímulos deben asignarse al azar.

Tipos de diseños cuasiexperimentales

Con excepción de la diferencia que acabamos de mencionar, los cuasiexperimentos son muy parecidos a los experimentos “puros”. Por tanto, podemos decir que *hay casi tantos diseños cuasiexperimentales como experimentales “puros”*. Sólo que no hay asignación al azar ni emparejamiento. Pero por lo demás son iguales, la interpretación es similar, las comparaciones son las mismas y los análisis estadísticos iguales (salvo que a veces se consideran las pruebas para datos no correlacionados). Es por ello que nos limitaremos a ver sólo algunos de los diseños cuasiexperimentales (el resto puede ser deducido de sus correspondientes diseños experimentales “puros”, además de quitarles la “R” de asignación al azar) y se comentarán brevemente porque las comparaciones, interpretaciones y los análisis son prácticamente equiparables.² Consideramos que no sería adecuado volver a explicar dichas comparaciones, interpretaciones y análisis.

1. Diseño con posprueba únicamente y grupos intactos.

Este primer diseño utiliza dos grupos: uno recibe el tratamiento experimental y el otro no. Los grupos son comparados en la posprueba para analizar si el tratamiento experimental tuvo un efecto sobre la variable dependiente (O_1 con O_2). El diseño

² Si al lector le surge alguna duda respecto a qué comparaciones, interpretaciones y análisis pueden llevarse a cabo en un diseño cuasiexperimental, le recomendamos revise el diseño experimental “puro” correspondiente, vale recordar que la diferencia es que en el cuasiexperimental no hay aleatorización y los grupos pueden estar no correlacionados.

puede diagramarse del siguiente modo, tal como se ve, no hay asignación al azar ni emparejamiento:

$$\begin{array}{ccc} G_1 & X & 0_1 \\ G_2 & - & 0_2 \end{array}$$

Observe que si los grupos no son equiparables entre sí, las diferencias en las pospruebas de ambos grupos se atribuirían a la variable independiente, pero también a otras razones diferentes, lo peor es que el investigador quizá no se dé cuenta de ello. Por ejemplo, supongamos que se lleva a cabo un cuasiexperimento para analizar el efecto de la retroalimentación que los médicos dan a sus pacientes (respecto a su conducta en el tratamiento prescrito) sobre la obediencia o apego al tratamiento. Se podría partir de la siguiente hipótesis: “los pacientes que reciban mayor realimentación de parte de sus médicos, acerca de cómo se comportan en el tratamiento prescrito, se apegarán más a dicho tratamiento”. Es decir, los médicos que informen más a sus pacientes sobre su conducta en el tratamiento prescrito propiciarán en los pacientes un mayor deseo de seguir el tratamiento. La cuestión es motivar al paciente. Entonces, el investigador toma dos grupos de pacientes. Un grupo recibe realimentación sobre su conducta en el tratamiento prescrito y el otro no. Posteriormente se evalúa qué tanto se apega cada grupo, en lo sucesivo, al tratamiento. Supongamos que obtenemos el siguiente resultado: $0_1 > 0_2$ (el grupo experimental se apega más al tratamiento), entonces deducimos que la hipótesis fue confirmada. Pero para establecer lo anterior, debemos analizar con mucho cuidado que sea posible comparar a los grupos.

Imaginemos que el grupo experimental estaba formado por pacientes que asisten a un hospital donde con frecuencia se dan pláticas motivadoras para que los enfermos sigan los tratamientos prescritos, mientras que el grupo de control estaba integrado por pacientes que asisten a un hospital donde no se le asigna importancia a ello. ¿A qué razón se le podrían atribuir con certeza los resultados: a la manipulación de la variable independiente, a que los grupos de pacientes provienen de diferentes hospitales, a ambos factores, a algún otro? Como los grupos no son razonablemente equiparables, no tendríamos la certeza de cuál fue la causa o qué tanto contribuyeron los diversos factores involucrados. Hay un problema de validez interna.

También podría ser que el grupo experimental estuviera compuesto por pacientes que, desde antes del experimento, tuvieran una motivación elevada para apegarse a tratamientos médicos; o tal vez actúen otros factores que provocarán diferencias iniciales entre los grupos. Por ello, es importante que los grupos sean inicialmente comparables y que durante el experimento no ocurra algo que los haga diferentes, con excepción de la presencia-ausencia del tratamiento experimental (por ejemplo, misma enfermedad y tratamiento médico, hospital, galeno que los atiende, instrucciones y lugar, equivalencia como grupos en género, edad, avance de la enfermedad, etc.; nada más imaginemos que el grupo experimental, en promedio, está “más enfermo” que el de control, y los pacientes lo saben; llega a suceder que los más enfermos se apeguen más al tratamiento). El criterio de los experimentos “puros” en relación con mantener la igualdad de los grupos (salvo la manipulación de la variable independiente) se aplica por igual a los cuasiexperimentos.

Puede extenderse el diseño para incluir más de dos grupos. Se tienen así diferentes tratamientos experimentales o niveles de manipulación. Su formato general sería:

G_1	X_1	0_1
G_2	X_2	0_2
G_3	X_3	0_3
•	•	•
•	•	•
•	•	•
G_k	X_k	0_k
G_{k+1}	—	0_{k+1}

El último grupo es de control.

Un ejemplo de este diseño sería con cuatro grupos escolares de un mismo semestre y carrera (licenciatura) en una universidad, como grupos del cuasiexperimento. Veámoslo esquemáticamente en la tabla 5.7.

Tabla 5.7 Diagrama de un ejemplo del diseño cuasiexperimental con posprueba únicamente

Universidad del Centro Escuela de Psicología Tercer semestre		
Grupo	X_1	O_1
Grupo	X_2	O_2
Grupo	X_3	O_3
Grupo	—	O_4

Recuerde que los grupos son intactos (no se generan) y ya se habían constituido por motivos diferentes al cuasiexperimento (en este caso, la elección de estudiar una carrera y la asignación de alumnos a los grupos por parte de la Escuela de Psicología). Los tratamientos experimentales podrían ser métodos educativos.

2. Diseño con prueba-posprueba y grupos intactos (uno de ellos de control).

Este diseño es similar al que incluye posprueba únicamente y grupos intactos, sólo que en este caso a los grupos se les administra una preprueba, la cual puede servir para verificar la equivalencia inicial de los grupos (si son equiparables no debe haber diferencias significativas entre las prepruebas de los grupos). Su esquema más sencillo sería el siguiente:

G_1	O_1	X	O_2
G_2	O_3	—	O_4

Aunque puede extenderse a más de dos grupos (niveles o modalidades de manipulación de la variable independiente), lo cual se esquematizaría así:

G_1	01	X_1	02
G_2	03	X_2	04
G_3	05	X_3	06
•	•	•	•
•	•	•	•
•	•	•	•
G_k	02k-1	X_k	02k
G_{k+1}	02k+1	—	02(k+1)

Las posibles comparaciones entre las mediciones de la variable dependiente y las interpretaciones son las mismas que en el diseño experimental de preprueba-posprueba con grupo de control; solamente que, en este segundo diseño cuasiexperimental, los grupos son intactos y en la interpretación de resultados debemos tomarlo en cuenta. Recuérdese todo lo que se ha dicho de la probable no equivalencia de los grupos. Este aspecto se aplica a todos los diseños cuasiexperimentales.

3. Diseños cuasiexperimentales de series cronológicas.

En ocasiones el investigador pretende analizar *efectos al mediano y largo plazos o los efectos de la administración del tratamiento experimental varias veces; pero no cuenta con la posibilidad de asignar al azar a los sujetos o participantes a los grupos del experimento*. En estos casos, pueden utilizarse los *diseños cuasiexperimentales*, salvo que los grupos deben ser intactos. En ambas situaciones se aplican mediciones repetidas de la variable dependiente y se inserta el tratamiento experimental entre dos de esas mediciones en, al menos, un grupo; mientras que a otro grupo no se le aplica ningún tratamiento en el periodo de “experimentación”. Aunque desde la

perspectiva de la literatura clásica sobre experimentos (véase Campbell y Stanley, 1966) se reconoce como cuasiexperimento a un diseño que no tiene grupo de control. Bien, hablemos brevemente de estos diseños.

Series cronológicas de un solo grupo

A un único grupo se le administran varias prepruebas, después se le aplica el tratamiento experimental y finalmente varias pospruebas. El diseño se diagrama así:

G 0₁ 0₂ 0₃ X 0₄ 0₅ 0₆

El número de mediciones está sujeto a las necesidades específicas de la investigación que realizamos.

Un ejemplo muy difundido de este diseño lo constituyó la evaluación de un programa que tenía por objeto disminuir la velocidad en carreteras del estado de Connecticut, Estados Unidos. (Campbell, 1975). Los investigadores recolectaron informes y datos de accidentes de tránsito correspondientes a varios años anteriores y ulteriores a la implantación del programa. Encontraron que después del programa el número de accidentes disminuyó. Pero como las distintas mediciones habían mostrado una pauta ascendente y descendente inestable durante varios años, no se podía tener la certeza de que el programa hubiese sido la razón del descenso en el número de accidentes (Weiss, 1990). Entonces, fue necesario comparar las estadísticas de Connecticut con las de otros cuatro estados vecinos en los que no se habían efectuado los cambios en los reglamentos de tránsito propuestos por el programa del mencionado estado. Estos otros cuatro estados actuaron como grupos de control. Finalmente, se observó que en los otros estados no se había registrado una disminución equivalente al número de accidentes. Las comparaciones dieron pie para concluir que el programa había generado los efectos deseados (Campbell, 1975; Glass, 1968).

Esta investigación cuasiexperimental en su primera etapa utilizó las series cronológicas de un solo grupo (figura 5.14).

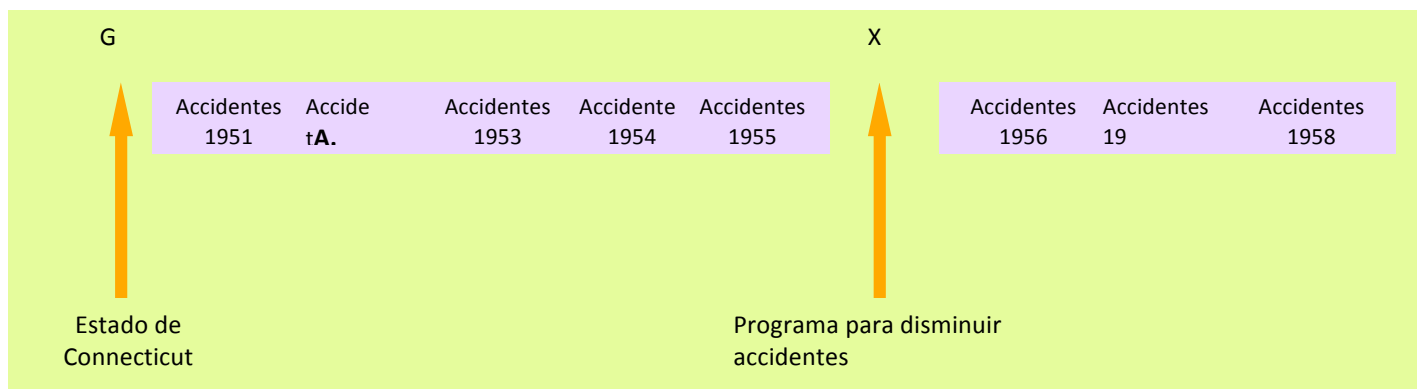
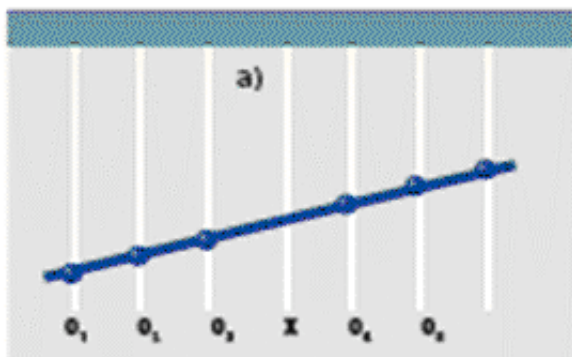


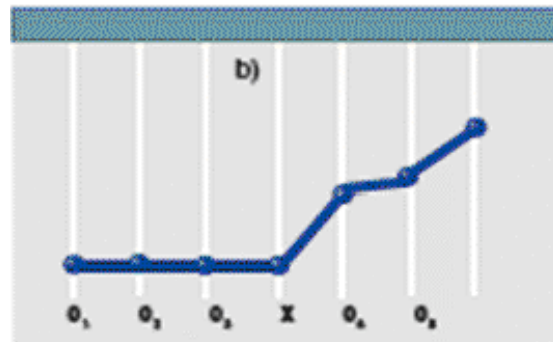
Figura 5.14 Ejemplo de un cuasiexperimento en su primera etapa

Escala de la variable dependiente



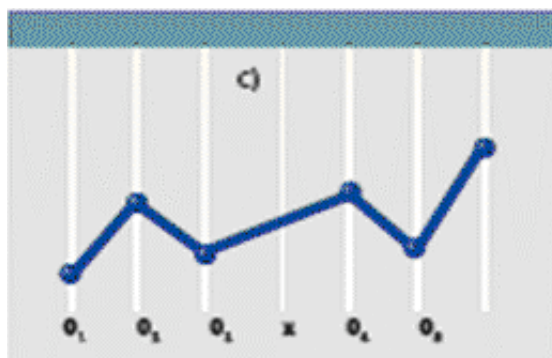
Interpretación: Aparentemente no hay efecto, sino una tendencia de aumento en la variable dependiente, que comienza desde antes del tratamiento.

Escala de la variable dependiente



Interpretación: posible efecto del tratamiento sobre la variable dependiente.

Escala de la variable dependiente



Interpretación: aunque entre la primera y la última mediciones se aprecian diferencias, no puede concluirse que haya o no efecto, pues el patrón es irregular.

Figura 5.15 Ejemplos de resultados en series cronológicas de un solo grupo

Otro ejemplo de este diseño sería medir las ventas de un producto durante varios meses, introducir una campaña publicitaria para ese producto y, después, medir durante meses el nivel de ventas.

Las series cronológicas de un solo grupo llegan a producir diversos *patrones de resultados*. A manera de ejemplo podríamos tener los patrones mostrados en la figura 5.15 (algunos de los cuales fueron expuestos en las series cronológicas experimentales).

En las series cronológicas de un único grupo debe tomarse muy en cuenta que no se tiene punto de comparación (grupo de control); por lo tanto, la interpretación del patrón en la variable dependiente (o patrones de las variables dependientes) tiene que ser muy cuidadosa, y habrá de analizarse si no han actuado o interactuado otras posibles causas, además del tratamiento experimental o variable independiente. La historia y el hecho de que el grupo sea atípico son riesgos que se afrontan en este diseño, al igual que la instrumentación. Normalmente, este diseño cuasiexperimental se utiliza con propósitos correlacionales y no explicativos.

Series cronológicas cuasiexperimentales con múltiples grupos.

Estos diseños pueden adoptar la estructura de las series cronológicas experimentales, con la diferencia de que en estas últimas los individuos se asignan al azar a los grupos, mientras que en las *cuasiexperimentales* tenemos *grupos intactos*. Por tanto, ocurrirían las mismas variaciones³ que se muestran en la tabla 5.8.

³ Recomendamos al lector que revise los apartados relativos a las series cronológicas experimentales antes de leer este apartado. De hecho, podrá notar que los diseños son los mismos, salvo que en los esquemas y diagramas de las series experimentales aparece el símbolo “R” de asignación al azar. Por ello, aquí se omiten explicaciones, interpretaciones y ejemplos; de lo contrario, pecaríamos de redundantes. Únicamente nos limitaremos a esquematizar los diseños, sin explicaciones, porque lo que puede decirse es lo mismo que lo dicho en las series cronológicas experimentales. Desde luego, debe resaltarse que en las series cuasiexperimentales los grupos son intactos y es necesario asegurar que los grupos sean equiparables.

Tabla 5.8 Ejemplos de diseños cuasiexperimentales con series cronológicas

Sin prepruebas y grupo de control

G ₁	X ₁	O ₁	O ₂	O ₃
G ₂	X ₂	O ₄	O ₅	O ₆
G ₃	X ₃	O ₇	O ₈	O ₉
G ₄	—	O ₁₀	O ₁₁	O ₁₂

Con prepruebas y grupo de control

G ₁	O ₁	O ₂	O ₃	X	O ₄	O ₅	O ₆
G ₂	O ₇	O ₈	O ₉	—	O ₁₀	O ₁₁	O ₁₂

Series cronológicas cuasiexperimentales con repetición de estímulo.

Estas series también son similares a sus correspondientes experimentales, pero con grupos intactos. Así, tendríamos los siguientes diagramas para ilustrarlas:

G ₁	O ₁	O ₂	X ₁	O ₃	O ₄	X ₁	O ₅	O ₆	O ₇	X ₁	O ₈	O ₉	O ₁₀
G ₂	O ₁₁	O ₁₂	—	O ₁₃	O ₁₄	—	O ₁₅	O ₁₆	O ₁₇	—	O ₁₈	O ₁₉	O ₂₀

G ₁	O ₁	X ₁	O ₂	X ₁	O ₃	X ₁	O ₄	X ₁	O ₅	O ₆
G ₂	O ₇	X ₂	O ₈	X ₂	O ₉	X ₂	O ₁₀	X ₂	O ₁₁	O ₁₂
G ₃	O ₁₃	—	O ₁₄	—	O ₁₅	—	O ₁₆	—	O ₁₇	O ₁₈

Series cronológicas cuasiexperimentales con tratamientos múltiples

Al igual que en los casos anteriores, estas series son similares a sus correspondientes experimentales, sólo que con grupos intactos. Por lo tanto, tendríamos diagramas como éstos:

G ₁	X ₁	0 ₁	0 ₂	X ₂	0 ₃	0 ₄	X ₃	0 ₅	0 ₆	0 ₇
G ₂	X ₂	0 ₈	0 ₉	X ₁	0 ₁₀	0 ₁₁	X ₃	0 ₁₂	0 ₁₃	0 ₁₄
G ₃	X ₃	0 ₁₅	0 ₁₆	X ₂	0 ₁₇	0 ₁₈	X ₁	0 ₁₉	0 ₂₀	0 ₂₁
G ₄	X ₂	0 ₂₂	0 ₂₃	X ₃	0 ₂₄	0 ₂₅	X ₁	0 ₂₆	0 ₂₇	0 ₂₈
G ₅	X ₁	0 ₂₉	0 ₃₀	X ₃	0 ₃₁	0 ₃₂	X ₂	0 ₃₃	0 ₃₄	0 ₃₅
G ₆	X ₃	0 ₃₆	0 ₃₇	X ₁	0 ₃₈	0 ₃₉	X ₂	0 ₄₀	0 ₄₁	0 ₄₂

TIPOS DE VARIABLES EN EXPERIMENTOS Y CUASIEXPERIMENTOS

Como complemento a lo que se menciona en el libro sobre variables y experimentos, únicamente queremos ahondar un poco sobre los tipos de variables en los diseños experimentales y cuasiexperimentales. Los principales tipos de variables son:

1. Independiente: tratamiento experimental que provoca efectos (causa).
2. Dependiente: efecto o consecuencia (provocado o provocada por el tratamiento o variable independiente).
3. Interviniente: moderador de la relación causal entre la variable independiente y dependiente. Si no se conoce su efecto o no se controla, el experimento puede invalidarse.
4. Explicaciones rivales o fuentes de invalidación interna (pueden ser variables independientes o intervinientes): su influencia debe conocerse o controlarse, de no ser así, el experimento puede invalidarse. Asimismo, es factible que se combine con otras variables para afectar a la dependiente.
5. Variable de control: influye en la dependiente, pero es neutralizada por el diseño o por los procedimientos estadísticos.

La diferencia entre la variable de control e interviniente reside en que en la primera se neutralizan sus efectos, y en la segunda se conocen éstos (Creswell, 2009). Un ejemplo sería el que se muestra en la figura 5.16

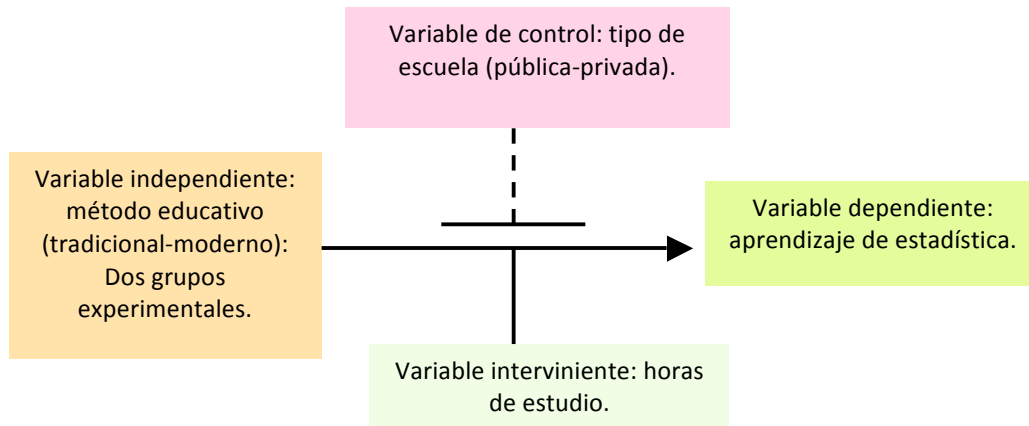


Figura 5.16 Diferencia entre variable interviniente y de control

La variable independiente se manipula (a un grupo se le expone a un método, al otro a un método distinto). El tipo de escuela se controla al asignar a los dos grupos por igual (en la misma proporción) alumnos de escuelas públicas y privadas (la composición de cada uno sería: 50% de estudiantes de instituciones públicas y 50% de estudiantes de escuelas privadas). Los efectos de la variable interviniente se conocen, al medir el número de horas dedicadas al estudio (con su introducción al análisis). La variable dependiente se mide. Los análisis estadísticos ayudan a esclarecer las relaciones entre todas las variables.

Nota final: Los pasos para un diseño cuasiexperimental son los mismos que para uno experimental.