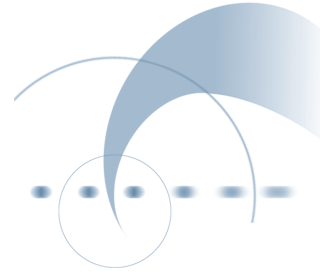
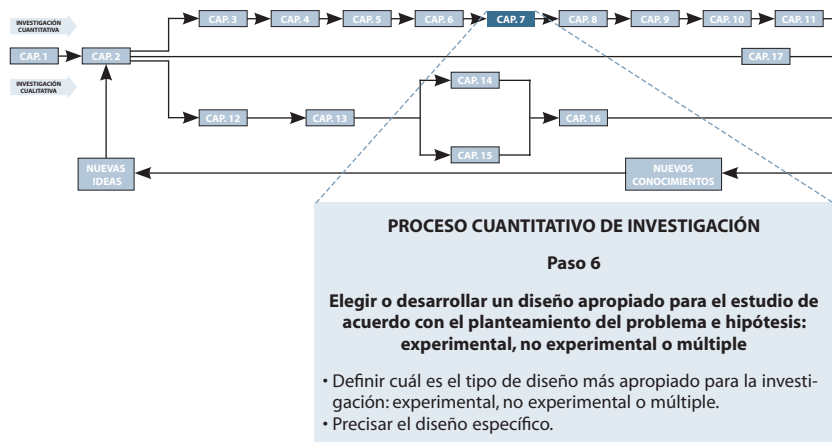


# Capítulo 7



## Concepción o elección del diseño de investigación



### OBJETIVOS DE APRENDIZAJE

Al terminar este capítulo, el alumno será capaz de:

- Definir el significado del término “diseño de investigación”, así como las implicaciones que se derivan de elegir uno u otro tipo de diseño.
- Comprender que en un estudio pueden incluirse uno o varios diseños de investigación.
- Conocer los tipos de diseños de la investigación cuantitativa y relacionarlos con los alcances del estudio.
- Comprender las diferencias entre la investigación experimental y la investigación no experimental.
- Analizar los diferentes diseños experimentales y sus grados de validez.
- Analizar los distintos diseños no experimentales y las posibilidades de investigación que ofrece cada uno.
- Realizar experimentos y estudios no experimentales.
- Comprender cómo el factor tiempo altera la naturaleza de un estudio.

### Síntesis

Con el propósito de responder a las preguntas de investigación planteadas y cumplir con los objetivos del estudio, el investigador debe seleccionar o desarrollar un diseño de investigación específico. Cuando se establecen y formulan hipótesis, los diseños sirven también para someterlas a prueba. Los diseños cuantitativos pueden ser experimentales o no experimentales.

En este capítulo se analizan diferentes diseños experimentales y la manera de aplicarlos. Asimismo, se explica el concepto de validez experimental y cómo lograrla.

También se presenta una clasificación de diseños no experimentales, en la que se considera: a) el factor tiempo o número de veces en que se recolectan datos y b) el alcance del estudio.

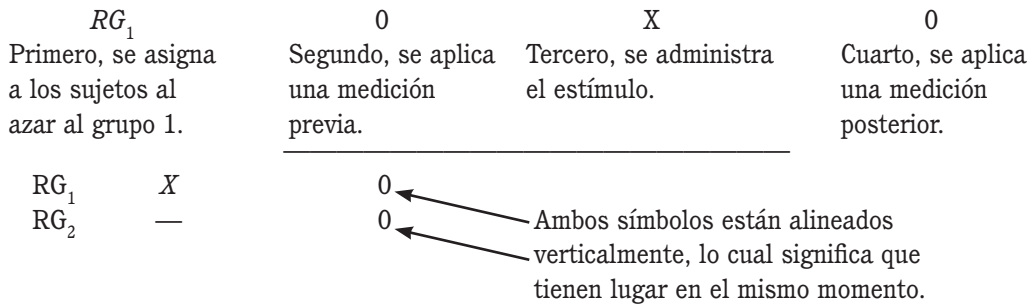
Del mismo modo, se deja en claro que ningún tipo de diseño es intrínsecamente mejor que otro, sino que son el planteamiento del problema, los alcances de la investigación y la formulación o no de hipótesis y su tipo, los que determinan qué diseño es el más adecuado para un estudio específico; asimismo, es posible utilizar más de un diseño.

efecto. Se trata de explicaciones rivales frente a la explicación de que las variables independientes afectan a las dependientes. En un libro clásico de Campbell y Stanley (1966) se definieron estas explicaciones rivales, las cuales han sido ampliadas y a las que se han agregado otras en referencias más recientes (por ejemplo, Campbell, 1975; Matheson, Bruce y Beauchamp, 1985; Christensen, 2000; Babbie, 2001; Creswell, 2005; Mertens, 2005).

A estas explicaciones se les conoce como fuentes de invalidación interna porque precisamente atentan contra la validez interna de un experimento. Ésta se refiere a cuánta confianza tenemos en que sea posible interpretar los resultados del experimento y éstos sean válidos. La **validez interna** se relaciona con la calidad del experimento y se logra cuando hay control, cuando los grupos difieren entre sí solamente en la exposición a la variable independiente (ausencia-presencia o en grados o modalidades), cuando las mediciones de la variable dependiente son confiables y válidas, y cuando el análisis es el adecuado para el tipo de datos que estamos manejando. El **control** en un experimento se alcanza eliminando esas explicaciones rivales o fuentes de invalidación interna. A continuación se mencionan y definen de acuerdo con Campbell y Stanley (1966), Campbell (1975), Babbie (2001), Creswell (2005) y Mertens (2005).

**Explicaciones rivales** Son las fuentes de invalidación interna, el control de un experimento se alcanza al ser eliminadas éstas.

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, si durante un experimento para analizar el efecto que distintas formas de retroalimentación en el trabajo tienen en la motivación, considerando dos grupos de obreros, a uno le aumentan el salario o se le reúne y 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.* Poca o nula confiabilidad de las mediciones, fluctuaciones en las personas seleccionadas o en los componentes del experimento, o inestabilidad autónoma de mediciones repetidas aparentemente “equivalentes”. 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. Además, 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 podamos estar seguros de ello. No tenemos confianza en los resultados.



## Preexperimentos

Los **preexperimentos** se llaman así porque su grado de control es mínimo.

### 1. Estudio de caso con una sola medición

Este diseño podría diagramarse de la siguiente manera:



Consiste en administrar un estímulo o tratamiento a un grupo y después aplicar una medición de una o más variables para observar cuál es el nivel del grupo en estas variables.

Este diseño no cumple con los requisitos de un experimento “puro”. No hay manipulación de la variable independiente (no hay varios niveles de ella, ni siquiera los niveles mínimos de presencia-ausencia). Tampoco hay una referencia previa de cuál era el nivel que tenía el grupo en la(s) variable(s) dependiente(s) antes del estímulo, ni existe grupo de comparación. El diseño adolece de los defectos que fueron mencionados al hablar de uno de los requisitos para lograr el control experimental: tener varios grupos de comparación. No es posible establecer causalidad con certeza ni se controlan las fuentes de invalidación interna.

### 2. Diseño de preprueba-posprueba con un solo grupo

Este segundo diseño se diagramaría así:



A un grupo se le aplica una prueba previa al estímulo o tratamiento experimental, después se le administra el tratamiento y finalmente se le aplica una prueba posterior al estímulo.

Este diseño ofrece una ventaja sobre el anterior: existe un punto de referencia inicial para ver qué nivel tenía el grupo en la(s) variable(s) dependiente(s) antes del estímulo. Es decir, hay un seguimiento del grupo. Sin embargo, el diseño no resulta conveniente para fines de establecer causalidad: no hay manipulación ni grupo de comparación, y es posible que actúen varias fuentes de invalidación interna, por ejemplo, la historia. Entre  $0_1$  y  $0_2$

**Preexperimento** Diseño de un solo grupo cuyo grado de control es mínimo. Generalmente es útil como un primer acercamiento al problema de investigación en la realidad.

## 1. Diseño con posprueba únicamente y grupo de control

Este diseño incluye dos grupos, uno recibe el tratamiento experimental y el otro no (grupo de control). Es decir, la manipulación de la variable independiente alcanza sólo dos niveles: presencia y ausencia. Los sujetos se asignan a los grupos de manera aleatoria. Cuando concluye la manipulación, a ambos grupos se les administra una medición sobre la variable dependiente en estudio.

El diseño se diagrama de la siguiente manera:

$RG_1$	$X$	$O_1$
$RG_2$	–	$O_2$

En este diseño, la única diferencia entre los grupos debe ser la presencia-ausencia de la variable independiente. Inicialmente son equivalentes y para asegurarse de que durante el experimento continúen siéndolo (salvo por la presencia o ausencia de dicha manipulación) el experimentador debe observar que no ocurra algo que sólo afecte a un grupo. La hora en que se efectúa el experimento debe ser la misma para ambos grupos (o ir mezclando un sujeto de un grupo con un sujeto del otro grupo, cuando la participación es individual), al igual que las condiciones ambientales y demás factores mencionados al hablar sobre la equivalencia de los grupos.

Wiersma y Jurs (2005) comentan que, de preferencia, la posprueba debe administrarse inmediatamente después de que concluya el experimento, en especial si la variable dependiente tiende a cambiar con el paso del tiempo. La posprueba se aplica de manera simultánea a ambos grupos.

La comparación entre las pospruebas de ambos grupos ( $O_1$  y  $O_2$ ) nos indica si hubo o no efecto de la manipulación. Si ambas difieren significativamente<sup>13</sup> ( $O_1 \neq O_2$ ), esto nos indica que el tratamiento experimental tuvo un efecto a considerar. Por lo tanto, se acepta la hipótesis de diferencia de grupos. Si no hay diferencias ( $O_1 = O_2$ ), ello indica que no hubo un efecto significativo del tratamiento experimental ( $X$ ). En este caso se acepta la hipótesis nula.

En ocasiones se espera que  $O_1$  sea mayor que  $O_2$ . Por ejemplo, si el tratamiento experimental es un método educativo que facilita la autonomía por parte del alumno, y si el investigador formula la hipótesis de que incrementa el aprendizaje, cabe esperar que el nivel de aprendizaje del grupo experimental, expuesto a la autonomía, sea mayor que el nivel de aprendizaje del grupo de control, no expuesto a la autonomía:  $O_1 > O_2$ .

En otras ocasiones se espera que  $O_1$  sea menor que  $O_2$ . Por ejemplo, si el tratamiento experimental es un programa de televisión que supuestamente disminuye el prejuicio, el nivel de éste

**Experimentos “puros”**  
Son aquellos que logran el control y la validez interna al reunir dos requisitos: 1) grupos de comparación y 2) equivalencia de los grupos.

<sup>13</sup> Los estudiantes frecuentemente se preguntan: ¿qué es una diferencia significativa? Si el promedio en la posprueba de un grupo en alguna variable es de 10 (por ejemplo), y en el otro es de 12, ¿esta diferencia es o no significativa? ¿Puede o no decirse que el tratamiento tuvo un efecto sobre la variable dependiente? A este respecto, cabe mencionar que existen pruebas o métodos estadísticos que nos indican si una diferencia entre dos o más cifras (promedios, porcentajes, puntuaciones totales, etc.) es o no significativa. Estas pruebas toman en cuenta aspectos como el tamaño de los grupos cuyos valores se comparan, las diferencias entre quienes integran los grupos y otros factores. Cada comparación entre grupos es distinta y ello lo consideran los métodos, los cuales se explicarán en el capítulo correspondiente al análisis cuantitativo de los datos. No resultaría conveniente exponerlos aquí, porque habría que clarificar algunos aspectos estadísticos en los cuales se basan tales métodos, lo que provocaría confusión, sobre todo entre quienes se inician en el estudio de la investigación.

## 2. Diseño con preprueba-posprueba y grupo de control

Este diseño incorpora la administración de prepruebas a los grupos que componen el experimento. Los sujetos se asignan al azar a los grupos, después a éstos se les aplica simultáneamente la preprueba; un grupo recibe el tratamiento experimental y otro no (es el grupo de control); por último, se les administra, también simultáneamente, una posprueba. El diseño se diagrama como sigue:

$RG_1$	$O_1$	$X$	$O_2$
$RG_2$	$O_3$	—	$O_4$

La adición de la preprueba ofrece dos ventajas: primera, las puntuaciones de las prepruebas sirven para fines de control en el experimento, pues al compararse las prepruebas de los grupos se evalúa qué tan adecuada fue la asignación aleatoria, lo cual es conveniente con grupos pequeños. En grupos grandes la técnica de distribución aleatoria funciona, pero cuando tenemos grupos de 15 personas o menos no está de más evaluar qué tanto funcionó la asignación al azar. La segunda ventaja reside en que es posible analizar el puntaje-ganancia de cada grupo (la diferencia entre las puntuaciones de la preprueba y la posprueba).

El diseño controla todas las fuentes de invalidación interna por las mismas razones que se argumentaron en el diseño anterior (diseño con posprueba únicamente y grupo de control). Y la administración de pruebas queda controlada, ya que si la preprueba afecta las puntuaciones de la posprueba lo hará de manera similar en ambos grupos, y se sigue cumpliendo con la esencia del control experimental. Lo que influye en un grupo deberá influir de la misma manera en el otro, para mantener la equivalencia entre ambos.

En algunos casos, para no repetir exactamente la misma prueba, se desarrollan dos versiones de la prueba que sean equivalentes (que produzcan los mismos resultados).<sup>17</sup> La historia se controla al observar que ningún acontecimiento afecte sólo a un grupo.

Es posible extender este diseño para incluir más de dos grupos, lo cual se diagramaría de una manera general del siguiente modo:

$RG_1$	$O_1$	$X_1$	$O_2$
$RG_2$	$O_3$	$X_2$	$O_4$
$RG_3$	$O_5$	$X_3$	$O_6$
•	•	•	•
•	•	•	•
•	•	•	•
$RG_k$	$O_{2k-1}$	$X_k$	$O_{2k}$
$RG_{k+1}$	$O_{2k+1}$	—	$O_{2(k+1)}$

<sup>17</sup> Hay procedimientos para obtener pruebas “paralelas” o “gemelas”, los cuales nos garantizan la equivalencia; se hablará de ellos en el capítulo sobre recolección de los datos en el proceso cuantitativo. Si no se utiliza un método que verdaderamente asegure la equivalencia de las pruebas, se corre el riesgo de que éstas no sean equiparables y entonces no se pueden comparar las puntuaciones producidas por ambas pruebas. Es decir, se pueden presentar las fuentes de invalidación interna: “inestabilidad”, “instrumentación” y “regresión estadística”.

Las posibles comparaciones en este diseño son: a) las prepruebas entre sí ( $O_1, O_3, O_5$  y  $O_7$ ), b) las pospruebas entre sí para analizar cuál fue el método de enseñanza más efectivo ( $O_2, O_4, O_6$  y  $O_8$ ), c) el puntaje-ganancia de cada grupo ( $O_1$  frente a  $O_2, O_3$  frente a  $O_4, O_5$  frente a  $O_6$  y  $O_7$  frente a  $O_8$ ), y d) los puntajes-ganancia de los grupos entre sí. Al igual que en todos los diseños experimentales, es posible tener más de una variable dependiente (por ejemplo, interés por los hábitos higiénicos, disfrute del método de enseñanza, etc.). En este caso, las prepruebas y pospruebas medirán diversas variables dependientes.

Veamos algunos posibles resultados de este ejemplo y sus interpretaciones:

1. Resultado:  $O_1 \neq O_2, O_3 \neq O_4, O_5 \neq O_6, O_7 \neq O_8$ ; pero  $O_2 \neq O_4, O_2 \neq O_6, O_4 \neq O_6$ .  
Interpretación: hay efectos de todos los tratamientos experimentales, pero son diferentes.
2. Resultado:  $O_1 = O_3 = O_5 = O_2 = O_6 = O_7 = O_8$ ; pero  $O_3 \neq O_4$ .  
Interpretación: no hay efectos de  $X_1$  ni  $X_3$ , pero sí hay efectos de  $X_2$ .
3. Resultado:  $O_1 = O_3 = O_5 = O_7$  y  $O_2 = O_4 = O_6 = O_8$ ; pero  $O_1, O_3, O_5$  y  $O_7 < O_2, O_4, O_6$  y  $O_8$ .  
Interpretación: no hay efectos de los tratamientos experimentales, sino un posible efecto de sensibilización de la preprueba o de maduración en todos los grupos (éste es parejo y se encuentra bajo control).

### 3. Diseño de cuatro grupos de Solomon

R. L. Solomon (1949) describió un diseño que era la mezcla de los dos anteriores (diseño con posprueba únicamente y grupo de control más diseño de preprueba-posprueba con grupo de control). La suma de estos dos diseños origina cuatro grupos: dos experimentales y dos de control, los primeros reciben el mismo tratamiento experimental y los segundos no reciben tratamiento. Sólo a uno de los grupos experimentales y a uno de los grupos de control se les administra la preprueba; a los cuatro grupos se les aplica la posprueba. Los participantes se asignan en forma aleatoria.

El diseño se diagrama así:

$RG_1$	$O_1$	$X$	$O_2$
$RG_2$	$O_3$	—	$O_4$
$RG_3$	—	$X$	$O_5$
$RG_4$	—	—	$O_6$

El diseño original incluye sólo cuatro grupos y un tratamiento experimental. Los efectos se determinan comparando las cuatro pospruebas. Los grupos uno y tres son experimentales, y los grupos dos y cuatro son de control.

La ventaja de este diseño es que el experimentador o la experimentadora tienen la posibilidad de verificar los posibles efectos de la preprueba sobre la posprueba, puesto que a unos grupos se les administra preprueba y a otros no. Es posible que la preprueba afecte la posprueba o que aquélla interactúe con el tratamiento experimental. Por ejemplo, con promedios de una variable determinada podría encontrarse lo que muestra la tabla 7.5.

Teóricamente  $O_2$  debería ser igual a  $O_5$ , porque ambos grupos recibieron el mismo tratamiento; asimismo,  $O_4$  y  $O_6$  deberían tener el mismo valor, porque ninguno recibió estímulo experimental. Pero  $O_2 \neq O_5$  y  $O_4 \neq O_6$ , ¿cuál es la única diferencia entre  $O_2$  y  $O_5$ , y entre  $O_4$  y  $O_6$ ? La respuesta

**Tabla 7.5** Ejemplo de efecto de preprueba en el diseño de Solomon

$RG_1$	$O_1 = 8$	X	$O_2 = 14$
$RG_2$	$O_3 = 8.1$	—	$O_4 = 11$
$RG_3$	—	X	$O_5 = 11$
$RG_4$	—	—	$O_6 = 8$

es la preprueba. Las diferencias pueden atribuirse a un efecto de la preprueba (la preprueba afecta, aproximadamente, tres puntos, y el tratamiento experimental también tres puntos, poco más o menos). Veámoslo de manera esquemática:

Ganancia con preprueba y tratamiento = 6

Ganancia con preprueba y sin tratamiento = 2.9 (casi 3).

Porque la técnica de distribución aleatoria hace al inicio equivalentes a los grupos, supuestamente la preprueba hubiera sido para todos cerca de ocho, si se hubiera aplicado a los cuatro grupos. La “supuesta ganancia” (supuesta porque no hubo preprueba) del tercer grupo, con tratamiento y sin preprueba, es de tres. Y la “supuesta ganancia” (supuesta porque tampoco hubo preprueba) del cuarto grupo es nula o inexistente [cero]).

Esto indica que cuando hay preprueba y estímulo se obtiene la máxima puntuación de 14, si sólo hay preprueba o estímulo la puntuación es de 11, y cuando no hay ni preprueba ni estímulo de ocho (calificación que todos deben tener inicialmente por efecto de la asignación al azar). También podría ocurrir un resultado como el de la tabla 7.6. En este caso, la preprueba no afecta (véase la comparación entre  $O_3$  y  $O_4$ ), y el estímulo sí (compárese  $O_5$  con  $O_6$ ); pero cuando el estímulo o tratamiento se junta con la preprueba se observa un efecto importante (compárese  $O_1$  con  $O_2$ ), un efecto de interacción entre el tratamiento y la preprueba.

El **diseño de Solomon** controla todas las fuentes de invalidación interna por las mismas razones que fueron explicadas desde el diseño con posprueba únicamente y grupo de control. La administración de pruebas se somete a un análisis minucioso. La historia la controla si se observa que ningún suceso afecte sólo a un grupo.

**Diseño de Solomon** Considera sólo cuatro grupos y un tratamiento experimental. Su ventaja radica en que se pueden verificar los posibles efectos de la preprueba sobre la posprueba.

**Tabla 7.6** Ejemplo del efecto de interacción entre la preprueba y el estímulo en el diseño de Solomon

$RG_1$	$O_1 = 7.9$	X	$O_2 = 14$
$RG_2$	$O_3 = 8$	—	$O_4 = 8.1$
$RG_3$	—	X	$O_5 = 11$
$RG_4$	—	—	$O_6 = 7.9$

## 4. Diseños experimentales de series cronológicas múltiples

Los tres diseños experimentales que se han comentado sirven más bien para analizar efectos inmediatos o a corto plazo. En ocasiones el experimentador está interesado 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, modelos de entrenamiento o estrategias de las psicoterapias. Asimismo, en otras situaciones se busca evaluar la evolución del efecto en el corto, mediano y largo plazos (no solamente el resultado). También, en ocasiones la aplicación del estímulo por una sola vez no tiene efectos (una dosis de un medicamento, un único programa televisivo, unos cuantos anuncios en la radio, etc.). En tales casos es conveniente adoptar diseños con varias pospruebas, o bien con diversas prepruebas y pospruebas, con repetición del estímulo, con varios tratamientos aplicados a un mismo grupo y otras condiciones. A estos diseños se les conoce como **series cronológicas experimentales** (véase capítulo 5 de CD anexo). En realidad el término “serie cronológica” se aplica a cualquier diseño que efectúe a través del tiempo varias observaciones o mediciones sobre una o más variables, sea o no experimental,<sup>18</sup> sólo que en este caso se les llama experimentales porque reúnen los requisitos para serlo.

**Serie cronológica** Diseño que efectúa a través del tiempo varias observaciones o mediciones sobre una o más variables, sea o no experimental (véase Capítulo 5 del CD anexo).

En estos diseños se pueden tener dos o más grupos y los participantes son asignados al azar. Los diseños experimentales de **series cronológicas múltiples** se exponen detalladamente en el capítulo “Otros diseños experimentales” del CD anexo.

## 5. Diseños factoriales

En ocasiones, el investigador o la investigadora pretenden analizar experimentalmente el efecto que sobre la(s) variable(s) dependiente(s) tiene la manipulación de más de una variable independiente. Por ejemplo, analizar el efecto que poseen sobre la productividad de los trabajadores: 1) la fuente de retroalimentación sobre el desempeño en el trabajo (vía el supervisor “cara a cara”, por escrito y por medio de los compañeros) y 2) el tipo de retroalimentación (positiva, negativa y positiva y negativa). En este caso se manipulan dos variables independientes. O bien, en otro ejemplo, determinar el efecto de tres medicamentos distintos (primera variable independiente, clase de medicamento) y la dosis diaria (segunda variable independiente, con dos niveles, supongamos 40 mg y 20 mg) sobre la cura de una enfermedad (variable dependiente). También aquí tenemos dos independientes. Pero podríamos tener tres o más: conocer cómo afectan en el nivel de aceleración de un vehículo (dependiente), el peso del chasis (dos diferentes pesos), el material

<sup>18</sup> En la terminología sobre el diseño suelen utilizarse dos términos: “series de tiempo (cronológicas)” y “estudios panel”. Markus (1979) dice que la diferencia principal entre ambas estriba en que las series de tiempo toman observaciones de un solo ente (individuo, país, empresa, etc.), un número de veces relativamente alto; mientras que en los estudios panel las observaciones se toman de varios entes, pero relativamente unas pocas veces (casi siempre cuatro o menos). Kessler y Greenberg (1981) coinciden con estas definiciones, y agregan que la recolección de observaciones de series de tiempo, en diversos entes o unidades de análisis, con frecuencia se conoce como series múltiples de tiempo o diseño panel de múltiples ondulaciones. En este libro, los autores hemos decidido usar el término “series cronológicas múltiples”.



pero sin preprueba (Campbell y Stanley, 1966). Babbie (2001) utiliza un excelente ejemplo de esta influencia: en un experimento diseñado para analizar si una película disminuye el prejuicio racial, la preprueba podría sensibilizar al grupo experimental y la película lograr un efecto mayor del que tendría si no se aplicara la preprueba (por ejemplo, si se pasara la película en un cine o en la televisión). Esto es, que la película sólo surtiría efecto cuando se administra la preprueba.

## 2. Efecto de interacción entre los errores de selección y el tratamiento experimental

Este factor se refiere a que se elijan personas con una o varias características que hagan que el tratamiento experimental produzca un efecto, que no se daría si las personas no tuvieran esas características. Por ejemplo, si seleccionamos trabajadores bastante motivados para un experimento sobre productividad, podría ocurrir que el tratamiento sólo tuviera efecto en este tipo de trabajadores y no en otros (únicamente funciona con individuos sumamente motivados). Ello se resolvería con una muestra representativa de todos los trabajadores o introduciendo un diseño factorial, y una de las variables fuera el grado de motivación (véanse diseños factoriales en el CD anexo, capítulo “Otros diseños experimentales”, capítulo 5).

A veces este factor se presenta cuando se reclutan voluntarios para la realización de algunos experimentos.

## 3. Efectos reactivos de los tratamientos experimentales

La “artificialidad” de las condiciones puede hacer que el contexto experimental resulte atípico, respecto a la manera en que se aplica regularmente el tratamiento (Campbell, 1975). Por ejemplo, a causa de la presencia de observadores y equipo, los participantes llegan a cambiar su conducta normal en la variable dependiente medida, la cual no se alteraría en una situación común donde se aplicara el tratamiento. Por ello, el experimentador tiene que ingeniárselas para hacer que los sujetos se olviden de que están en un experimento y no se sientan observados. A esta fuente también se le conoce como “efecto Hawthorne”, por una serie de experimentos muy famosos desarrollados —entre 1924 y 1927— en una planta del mismo nombre de la Western Electric Company, donde al variar las condiciones de iluminación se obtenían incrementos en la productividad de los trabajadores, pero por igual al aumentar la luz que al disminuirla y, más bien, los cambios en la productividad se debieron a que los participantes se sentían atendidos (Ballantyne, 2000; Mayo, 1933).

## 4. Interferencia de tratamientos múltiples

Si se aplican varios tratamientos a un grupo experimental para conocer sus efectos por separado y en conjunto (por ejemplo, en infantes enseñarles hábitos higiénicos con un video, más una dinámica que implique juegos, más un libro explicativo); incluso, si los tratamientos no son de efecto reversible, es decir, si no es posible borrar sus efectos, las conclusiones solamente podrán hacerse extensivas a los infantes que experimenten la misma secuencia de tratamientos, sean múltiples o la repetición del mismo (véanse diseños con diversos tratamientos en el capítulo “Otros diseños experimentales”, capítulo 5 del CD anexo).

control directo sobre dichas variables ni se puede influir sobre ellas, porque ya sucedieron, al igual que sus efectos.

La **investigación no experimental** es un parteaguas de un gran número de estudios cuantitativos como las encuestas de opinión (*surveys*), los estudios *ex post-facto* retrospectivos y prospectivos, etc. Para ilustrar la diferencia entre un estudio experimental y uno no experimental consideremos el siguiente ejemplo. Claro está que no sería ético un experimento que obligara a las personas a consumir una bebida que afecta gravemente la salud. El ejemplo es sólo para ilustrar lo expuesto y quizá parezca un tanto burdo, pero es ilustrativo.

---

## ● EJEMPLO

---

### Para esclarecer la diferencia entre la investigación experimental y la investigación no experimental

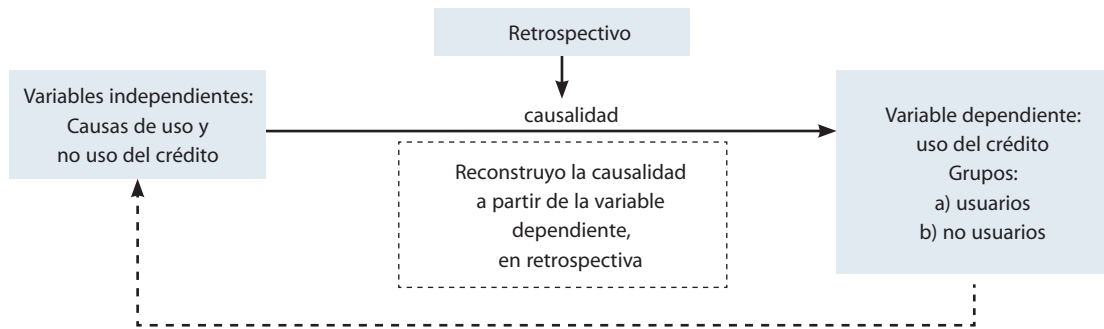
Vamos a suponer que un investigador desea analizar el efecto que produce el consumo de alcohol sobre los reflejos humanos. Su hipótesis es: a mayor consumo de alcohol, mayor lentitud en los reflejos de las personas. Si decidiera seguir un enfoque experimental, asignaría al azar los sujetos a varios grupos. Supónganse cuatro grupos: un primer grupo donde los participantes ingirieran un elevado volumen de alcohol (siete copas de tequila o brandy), un segundo grupo que tuviera un consumo medio de alcohol (cuatro copas), un tercer grupo que bebiera poco alcohol (una sola copa) y un cuarto grupo de control que no ingiriera nada de alcohol. Controlaría el lapso en el que todos los sujetos consumen su “ración” de alcohol (copa), así como otros factores (misma bebida, cantidad de alcohol servida en cada copa, etc.). Finalmente, mediría la calidad de la respuesta de los reflejos en cada grupo y compararía los grupos para determinar el efecto del consumo de alcohol sobre los reflejos humanos, y probar o desaprobar su hipótesis.

Desde luego, el enfoque podría ser cuasiexperimental (grupos intactos) o asignar los sujetos a los grupos por emparejamiento (digamos en cuanto al sexo, que influye en la resistencia al alcohol, pues la mayoría de las mujeres suelen tolerar menos cantidades que los hombres).

Por el contrario, si decidiera seguir un enfoque no experimental, el investigador podría acudir a lugares donde se localicen distintas personas con diferentes consumos de alcohol (por ejemplo, oficinas donde se haga la prueba del nivel de consumo de alcohol, como una estación de policía). Encontraría a personas que han bebido cantidades elevadas, medias y bajas de alcohol, así como a quienes no lo han ingerido. Mediría la calidad de sus reflejos, llevaría a cabo sus comparaciones y establecería el efecto del consumo de alcohol sobre los reflejos humanos, analizando si aporta evidencia en favor o en contra de su hipótesis.

---

Ahora, vayamos más a fondo para analizar las diferencias entre ambos tipos de investigación. En un **estudio experimental** se construye la situación y se manipula de manera intencional a la variable independiente (en este caso, el consumo del alcohol), después se observa el efecto de esta manipulación sobre la variable dependiente (aquí, la calidad de los reflejos). Es decir, el investigador influyó directamente en el grado de consumo de alcohol de los participantes. En la



**Figura 7.9** Ejemplo de una reconstrucción causal retrospectiva.

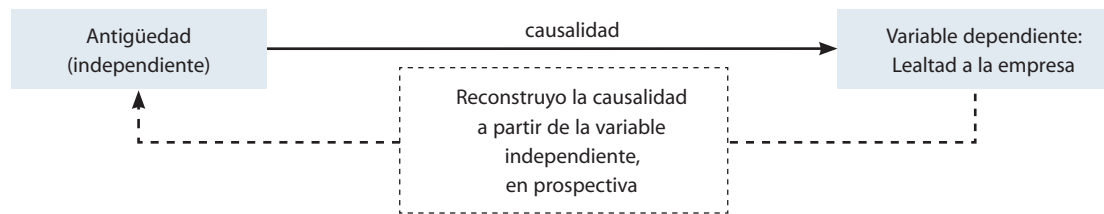
Veamos ahora una **investigación causal prospectiva**: Imaginemos que deseo indagar si la variable antigüedad provoca o no mayor lealtad a la empresa y por qué. Entonces, divido a los empleados en la variable independiente: *a*) muy alta antigüedad (25 o más años de laborar en la organización), *b*) alta antigüedad (16 a 24 años), *c*) mediana antigüedad (9 a 15 años), *d*) baja antigüedad (cuatro a ocho años), *e*) muy baja antigüedad (uno a tres años) y *f*) recién ingreso (un año o menos). Posteriormente, mido los niveles de lealtad y cuestiono a los empleados sobre cómo la antigüedad ha generado o no mayor lealtad. Así determino los efectos de interés. (Véase la figura 7.10.)

Para una ampliación de los diseños prospectivos y retrospectivos, se recomienda a León y Montero (2003).

En los diseños donde se reconstruyen las relaciones sobre la base de variabilidad amplia de las independientes y dependientes, no se parte de una variable en especial ni de grupos, sino que se evalúa la estructura causal completa (las relaciones en su conjunto).

Todos los estudios transeccionales causales nos brindan la oportunidad de predecir el comportamiento de una o más variables a partir de otras, una vez que se establece la causalidad. A estas últimas, se les denomina **variables predictoras**. Tales diseños requieren de análisis multivariados que se mencionan en el capítulo 8 del CD anexo (capítulo “Análisis estadístico-multivariado de los datos”). Aquí simplemente incluimos un ejemplo de una estructura causal compleja, que será evaluado en el capítulo mencionado. Lo importante es que se comprenda cómo en ocasiones se analizan múltiples variables y secuencias causales.

**Variables predictoras** Son aquellas que permiten, en los modelos transeccionales causales, predecir el comportamiento de una o más variables, una vez que se establece la causalidad.



**Figura 7.10** Ejemplo de una reconstrucción causal prospectiva.

## Diferencia entre diseños de tendencia y de evolución de grupo

La diferencia entre los diseños de evolución de grupo y los diseños de tendencia se observa en el siguiente ejemplo, adaptado de Wiersma (1999):

### ● EJEMPLO

#### Un ejemplo de un diseño de tendencia

Un investigador está interesado en estudiar las actitudes de los maestros respecto a las asociaciones de profesionales en la región “A”. Las actitudes se miden cada tres años durante un periodo de 15 años. En cada momento que se hace la medición, se selecciona, de la población de maestros existente en ese momento, una muestra de ellos. La membresía de la población puede cambiar a través del tiempo al menos parcialmente (algunos pueden dejar de ser maestros o ingresar nuevos maestros), pero en cualquier momento o tiempo la población es la misma: los maestros de la región “A” (llamada población general).

Si el investigador estuviera interesado en estudiar las actitudes hacia los sindicatos de profesionales por parte de los maestros que se iniciaron como tales en 2003 en la región “A” (cada cinco años), el estudio en ese año incluyó el análisis de una subpoblación o un grupo específico. Cinco años después, la siguiente muestra se obtendría de lo que queda de esa población, la cual en 2008 estaría constituida por maestros con cinco años de experiencia. Desde luego, algunos de los maestros que empezaron como tales en 2003 habrán dejado la docencia, y el estudio incluirá sólo las actitudes del grupo o la subpoblación de maestros que comenzaron a serlo en dicho año y que en 2008 continúan en el magisterio (de toda la población de maestros se estudia a una subpoblación). Éste sería un ejemplo de diseño de evolución de grupo o *cohort*.

En algunas poblaciones que se modifican con relativa facilidad, los miembros actuales de la población pueden cambiar totalmente a través del tiempo (Wiersma y Jurs, 2005). Por ejemplo, si se llevara a cabo una investigación sobre las opiniones políticas de estudiantes graduados en economía cada cinco años, habría un elevado porcentaje de cambio en los integrantes actuales de esa subpoblación. Aunque la subpoblación seguiría siendo siempre la misma: los graduados en economía de tal escuela. Es decir, los nombres de muchas personas cambiarían, pero la subpoblación no.

En los **diseños de tendencia y de evolución de grupo** se estudia el cambio en subpoblaciones o poblaciones pero, debido a que en cada momento o tiempo se recolectan datos sobre una muestra diferente aunque equivalente, el cambio se evalúa colectivamente y no de manera individual (porque las personas pueden cambiar). Si hay variaciones, el investigador no sería capaz de establecer en forma específica qué individuos las provocan. En ambos tipos de diseño tal situación se grafica como se muestra en la figura 7.14.

Es decir, es posible que algunos o todos los sujetos cambien, aunque la población o subpoblación es la misma.

**Diseños de tendencia y de evolución de grupo** Monitorean cambios en una población o subpoblación a través del tiempo, usando una serie de muestras que abarcan a grupos diferentes de participantes en cada ocasión.

**Tabla 7.8** Correspondencia entre tipos de estudio, hipótesis y diseño de investigación

Estudio	Hipótesis	Diseño
Exploratorio	— No se establecen, lo que se puede formular son conjeturas iniciales	— Transeccional descriptivo — Preexperimental
Descriptivo	— Descriptiva	— Preexperimental — Transeccional descriptivo
Correlacional	— Diferencia de grupos sin atribuir causalidad	— Cuasiexperimental — Transeccional correlacional — Longitudinal (no experimental)
	— Correlacional	— Cuasiexperimental — Transeccional correlacional — Longitudinal (no experimental)
Explicativo	— Diferencia de grupos atribuyendo causalidad	— Experimental — Cuasiexperimental, longitudinal y transeccional causal (cuando hay bases para inferir causalidad, un mínimo de control y análisis estadísticos apropiados para analizar relaciones causales)
	— Causales	— Experimental — Cuasiexperimental, longitudinal y transeccional causal (cuando hay bases para inferir causalidad, un mínimo de control y análisis estadísticos apropiados para analizar relaciones causales)

en cuatro grupos: uno donde se propicie una elevada motivación, otro con mediana motivación, otro más con baja motivación y un último al que no se le administre ningún motivador. Después compararíamos la productividad de los grupos. Tendríamos un experimento.

Si se tratara de grupos intactos tendríamos un cuasiexperimento. En cambio, si midiéramos la motivación existente en los trabajadores, así como su productividad y relacionáramos ambas variables, estaríamos realizando una investigación transeccional correlacional. Y si cada seis meses midiéramos las dos variables y estableciéramos su correlación efectuaríamos un estudio longitudinal.

## Los estudios de caso

Los **estudios de caso** son considerados por algunos autores y autoras como una clase de diseños, a la par de los experimentales, no experimentales y cualitativos (por ejemplo: Mertens, 2005 y Williams, Grinnell y Unrau, 2005), mientras que otros(as) los ubican como una clase de diseño experimental (v. gr., León y Montero, 2003) o un diseño etnográfico (Creswell, 2005). También han sido concebidos como un asunto de muestreo (Hernández, Fernández y Baptista, 2003) o un método (Yin, 2003).

## EJEMPLOS DESARROLLADOS

### LA TELEVISIÓN Y EL NIÑO

La investigación utilizará un diseño no experimental transversal correlacional-causal. Primero describirá el uso que los niños de la ciudad de México hacen de los medios de comunicación colectiva, el tiempo que dedican a ver la televisión, sus programas preferidos, las funciones y gratificaciones que la televisión tiene para los niños y otras cuestiones similares. Posteriormente, analizará los usos y las gratificaciones de la televisión en niños de diferentes niveles socioeconómicos, edades, géneros y otras variables (se relacionarán nivel socioeconómico y uso de la televisión, entre otras asociaciones).

Un caso de un estudio experimental sobre la televisión y el niño, consistiría en exponer durante determinado tiempo a un grupo de niños a tres horas diarias de televisión, otro a dos horas diarias, un tercero a una hora, y por último, un cuarto que no se expondría a la televisión. Todo ello para conocer el efecto que tiene la cantidad de horas expuestas ante contenidos televisivos (variable independiente) sobre diferentes variables dependientes (por ejemplo, autoestima, creatividad, socialización).

### EL CLIMA ORGANIZACIONAL

El estudio utilizará también un diseño no experimental transversal correlacional-causal. Asociará las dimensiones del clima organizacional: moral, apoyo de la dirección, innovación, percepción de la empresa-identidad-identificación, comunicación, percepción del desempeño, motivación intrínseca, autonomía, satisfacción general, liderazgo, visión y recompensas o retribución. Posteriormente buscará conocer su estructura causal.

### EL ABUSO SEXUAL INFANTIL

Se trata de un diseño experimental. Los datos se obtendrán de 150 preescolares

de tres centros de desarrollo infantil con una población similar, hijas e hijos de madres que laboran para la Secretaría de Educación del Estado de Querétaro. Se evaluarán seis grupos escolares que serán asignados a tres grupos experimentales. El primer grupo ( $n = 49$  niños) será evaluado al terminar un programa de prevención del abuso sexual infantil (PPASI); el segundo será medido después de un año de haber concluido el mismo programa (PPASI) (seguimiento,  $n = 22$  niños); y el tercero, un grupo de control que no será expuesto a algún PPASI particular ( $n = 79$  niños). A todos los integrantes de los grupos se les aplicarán tanto las escalas conductuales como la cognitiva. Las condiciones de recolección de datos seguirán el protocolo establecido por cada escala, en un espacio físico similar y de manera individual. La persona que evaluará será la misma en todos los casos, para evitar sesgos interobservadores. Es decir, se trata de un diseño experimental:

$G_1$	$X_1$	(evaluación inmediata al terminar el PPASI)	$O_1$
$G_2$	$X_2$	(evaluación a un año de concluir el PPASI)	$O_2$
$G_3$	—	(sin PPASI)	$O_3$

$O_1$ ,  $O_2$  y  $O_3$  son mediciones conductuales y cognitivas

Estímulo (PPASI) por medio del taller: "Porque me quiero, me cuido", se basará principalmente en la mejora de la autoestima, el manejo y expresión de sentimientos, la apropiación de su cuerpo, la discriminación de contactos apropiados e inapropiados, la asertividad, el esclarecimiento de redes de apoyo y prácticas para pedir ayuda denunciando el abuso. Las técnicas usadas en dicho taller principalmente serán: modelado, ensayo, cuento, retroalimentación, actuación y dibujo. El programa se llevará a cabo a lo largo del