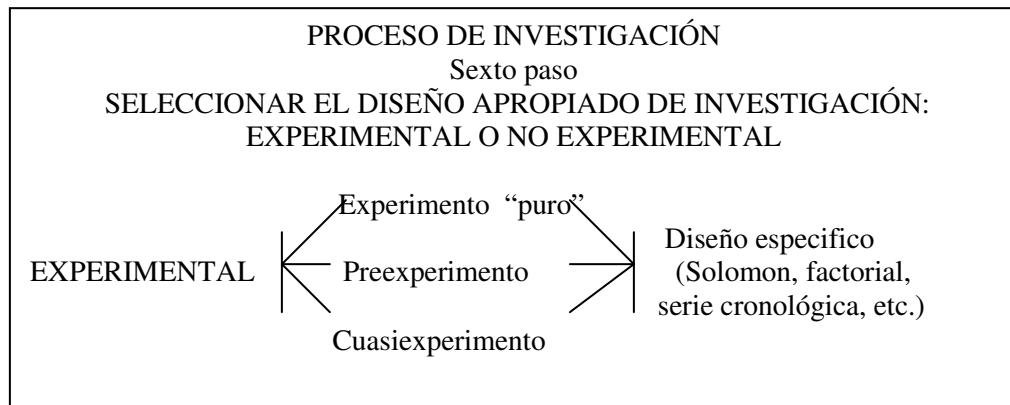


6

Diseños experimentales de investigación: preexperimentos, experimentos “verdaderos” y cuasiexperimentos



OBJETIVOS DE APRENDIZAJE

Que el alumno:

- 1) Comprenda las diferencias entre la investigación experimental y la investigación no experimental.
- 2) Analice los diferentes diseños experimentales y sus grados de validez.
- 3) Comprenda los conceptos de validez interna y validez externa.
- 4) Se encuentre habilitado para realizar experimentos válidos.
- 5) Esté capacitado para evaluar experimentos que hayan sido realizados.

SÍNTESIS

Con el propósito de responder a las preguntas de investigación planteadas y someter a prueba las hipótesis formuladas se selecciona un diseño específico de investigación. Los diseños pueden ser experimentales o no experimentales.

En este capítulo se analizan diferentes diseños experimentales y la manera de aplicarlos. Asimismo

se discute el concepto de validez experimental y cómo lograrla.

6.1. ¿QUÉ ES UN DISEÑO DE INVESTIGACIÓN?

Una vez que se ha definido el tipo de estudio a realizar y establecido la(s) hipótesis de investigación o los lineamientos para la investigación (si es que no se tienen hipótesis), el investigador debe concebir la manera práctica y concreta de responder a las preguntas de investigación. Esto implica seleccionar o desarrollar un *diseño de investigación* y aplicarlo al contexto particular de su estudio. *El término “diseño” se refiere al plan o estrategia concebida para responder a las preguntas de investigación* (Christensen, 1980). El diseño señala al investigador lo que debe hacer para alcanzar sus objetivos de estudio, contestar las interrogantes que se ha planteado y analizar la certeza de la(s) hipótesis formuladas en un contexto en particular. Por ejemplo, si la pregunta de investigación coloquial era: ¿Le gustará a Ana: Por qué sí y por qué no? y la hipótesis: “Yo le resulto atractivo a Ana porque así me lo ha hecho saber”.

El diseño sería el plan o la estrategia para confirmar si es o no cierto que le resulto atractivo a Ana (el plan incluiría actividades tendientes a encontrar la respuesta a la pregunta de investigación). En este caso podrían ser: “el día de mañana buscaré a Ana después de la clase de Estadística, me acercaré a ella, le diré que se ve muy guapa y la invitaré a tomar un café. Una vez que estemos en la cafetería la tomaré de la mano, y si ella no retira su mano, la invitaré a cenar el siguiente fin de semana y si acepta, en el lugar donde cenemos le diré que ella me resulta atractiva y le preguntaré si yo le resulto atractivo”. Desde luego, yo pude haber seleccionado o concebido otra estrategia, tal como invitarla a bailar o al cine en lugar de cenar; o bien si conozco a varias amigas de Ana y yo también soy muy amigo de ellas, preguntarles si le resulto atractivo a Ana. En el estudio del comportamiento humano disponemos de distintas clases de diseños o estrategias para poder investigar y debemos elegir un diseño entre las alternativas existentes.

Si el diseño está bien concebido, el producto último de un estudio (sus resultados) tendrá mayores posibilidades de ser válido (Kerlinger, 1979). Y no es lo mismo seleccionar un tipo de diseño que otro, cada uno tiene sus características propias —como se verá más adelante—. No es lo mismo preguntarle directamente a Ana si le resulto o no atractivo que preguntarle a sus amigas, o que en lugar de preguntarle verbalmente prefiero analizar su conducta no verbal (cómo me mira, qué reacciones tiene cuando la abrazo o me acerco a ella, etcétera). Como tampoco es igual si le pregunto en presencia de otras personas que si le pregunto solamente estando los dos. La precisión de la información obtenida puede variar en función del diseño o estrategia elegida.

6.2. ¿DE QUÉ TIPOS DE DISEÑOS DISPONEMOS PARA INVESTIGAR EL COMPORTAMIENTO HUMANO?

En la literatura sobre la investigación podemos encontrar diferentes clasificaciones de

los tipos de diseños existentes. En este libro se adoptará la siguiente clasificación: *investigación experimental e investigación no experimental*. A su vez, la *investigación experimental puede dividirse de acuerdo con las categorías de Campbell y Stanley (1966) en: preexperimentos, experimentos puros (verdaderos) y cuasiexperimentos*. La investigación no experimental será subdividida en diseños transaccionales o transversales y diseños longitudinales. Dentro de cada clasificación se comentarán diseños específicos.

Cabe aclarar que en términos generales, *los autores no consideramos que un tipo de investigación sea mejor que otro* (experimental versus no experimental). Como menciona Kerlinger (1979): ““Los dos tipos de investigación son relevantes y necesarios, tienen un valor propio y ambos deben llevarse a cabo”. Cada uno posee sus características y la elección sobre qué clase de investigación y diseño específico hemos de seleccionar, depende de los objetivos que nos hayamos trazado, las preguntas planteadas, el tipo de estudio a realizar (exploratorio, descriptivo, correlacional o explicativo) y las hipótesis formuladas.

6.3. ¿QUÉ ES UN EXPERIMENTO?

El término ‘experimento’ puede tener —al menos— dos acepciones, una general y otra particular. La general se refiere a “tomar una acción” y después observar las consecuencias de una acción (Babbie, 1979). Este uso del término es bastante coloquial, así hablamos —por ejemplo— de “experimentar” cuando mezclamos sustancias químicas y vemos la reacción de este hecho o cuando nos cambiamos de peinado y vemos el impacto que provoca en nuestros amigos esta transformación. La esencia de esta concepción de “experimento” es que éste involucra la manipulación intencional de una acción para analizar sus posibles efectos. La acepción particular que va más de acuerdo con un sentido científico del término, se refiere a “un estudio de investigación en el que se manipulan deliberadamente una o más variables independientes (supuestas causas) para analizar las consecuencias de esa manipulación sobre una o más variables dependientes (supuestos efectos), dentro de una situación de control para el investigador”. Esta definición puede parecer compleja, sin embargo, conforme se vayan analizando sus componentes se irá aclarando su sentido.

Los experimentos “auténticos o puros” manipulan variables independientes para ver sus efectos sobre variables dependientes en una situación de control.

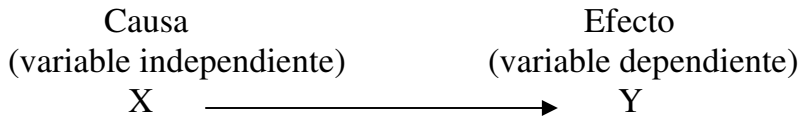
Si tomamos la acepción general del término “experimento”, los preexperimentos, los experimentos ““verdaderos” y los cuasiexperimentos podrían considerarse experimentos, ya que como se detallará más adelante ““toman una acción” y miden su efecto o efectos. En cambio, si tomamos la segunda acepción (que hemos llamado particular”), sólo los experimentos “puros” serían ““experimentos” y ambos conceptos se considerarían

equiparables. En este capítulo nos centraremos en los experimentos ““verdaderos o puros”.

6.4. ¿CUAL ES EL PRIMER REQUISITO DE UN EXPERIMENTO “PURO”?

El primer requisito de un experimento puro es la *manipulación intencional de una o más variables independientes*.

La variable independiente es la que se considera como supuesta *causa* en una relación entre variables, es la condición antecedente; y al *efecto* provocado por dicha causa se le denomina *variable dependiente* (consecuente).



Y como se mencionó en el capítulo anterior referente a las hipótesis, el investigador puede considerar dentro de su estudio a dos o más variables independientes. Cuando realmente existe una relación causal entre una variable independiente y una dependiente, al hacer variar intencionalmente a la primera, la segunda tendrá que variar. Si la motivación es causa de la productividad, al variar la motivación deberá variar la productividad.

Un experimento se lleva a cabo para analizar si una o más variables independientes afectan a una o más variables dependientes y por qué las afectan. Por ahora, simplifiquemos el problema de estudio a una variable independiente y una dependiente. En un auténtico experimento, la variable independiente resulta de interés para el investigador porque es la variable que se hipotetiza será una de las causas que producen el efecto supuesto (Christensen, 1980). Para obtener evidencia de esta relación causal supuesta, el investigador manipula la variable independiente para ver su efecto sobre la dependiente. Es decir, hace variar a la independiente y observa si la dependiente varía o no. Manipular es sinónimo de hacer variar o dar distintos valores a la variable independiente.

EJEMPLO

Si un investigador deseara analizar el posible efecto de los contenidos televisivos antisociales sobre la conducta agresiva de determinados niños podría hacer que un grupo viera un programa de televisión con contenido antisocial y otro grupo viera un programa con contenido prosocial,²¹ y posteriormente observaría cual de los dos grupos muestra una mayor conducta agresiva. La hipótesis de investigación nos hubiera señalado algo así:

‘La exposición por parte de los niños a contenidos antisociales tenderá a provocar un aumento en su conducta agresiva’. Si descubre que el grupo que vio el programa antisocial muestra una mayor conducta agresiva respecto al grupo que vio el programa prosocial, y descubre que no hay otra posible causa que hubiera afectado a los grupos de manera desigual; comprobaría su hipótesis.

²¹ En este momento no se discute sobre el método para asignar a los niños a los dos grupos, ello se discutirá en el apartado de control y validez interna. Lo que importa por ahora es que se comprenda el significado de manipulación de la variable independiente.

En el ejemplo, el investigador está manipulando o haciendo variar a la variable independiente para observar el efecto sobre la dependiente. ¿Cómo está manipulando la independiente? Lo hace dándole dos valores: presencia de contenidos antisociales en la televisión (programa antisocial) y ausencia de contenidos antisociales por televisión (programa prosocial). La variación es hecha a propósito por el experimentador (no es casual), éste tiene control directo sobre la manipulación, crea las condiciones para proveer el tipo de variación deseada. En un experimento, para que una variable pueda ser calificada como independiente se necesitan dos requerimientos: que varíe o sea manipulada y que esta variación pueda controlarse.

La Variable dependiente se mide

Por su parte, *la variable dependiente no se manipula, sino que se mide* para ver el efecto de la manipulación de la variable independiente sobre ella. Esto podría esquematizarse de la siguiente manera:

Manipulación de la variable independiente	Medición del efecto sobre la variable dependiente
X_A	Y
X_B	
.	
.	
.	

Las letras “A, B,...” indicarían distintos niveles de variación de la independiente.

Grados de manipulación de la variable independiente

La manipulación o variación de una variable independiente puede llevarse a cabo en dos o más grados. El nivel mínimo de manipulación es dos: presencia-ausencia de la variable independiente. Cada nivel o grado de manipulación implica un grupo en el experimento.

Presencia-ausencia

Implica que un grupo se expone a la presencia de la variable independiente y el otro no. Luego los dos grupos son comparados para ver si el grupo que se expuso a la variable independiente difiere del grupo no expuesto a ésta. En el ejemplo anterior del posible efecto

del contenido antisocial de la televisión sobre la conducta agresiva de ciertos niños, un grupo era expuesto a la variable independiente y el otro no. Al grupo que se expone a la presencia de la variable independiente se le conoce como “grupo experimental” y al grupo en el cual está ausente dicha variable se le denomina grupo de control”. Aunque en realidad ambos grupos participan en el experimento.

A la presencia de la variable independiente muy frecuentemente se le llama “tratamiento experimental” o “estímulo experimental”. Es decir, el grupo experimental recibe el tratamiento o estímulo experimental o lo que es lo mismo, se le expone a la variable independiente; mientras que el grupo de control no recibe el tratamiento o estímulo experimental. Otro ejemplo sería el siguiente: Supongamos que pretendemos investigar si un medicamento es o no útil para la cura de alguna enfermedad. Al grupo experimental se le administra el medicamento (presencia de la variable independiente o tratamiento experimental) y al grupo de control no, a este último se le administra un placebo (por ejemplo, un aparente medicamento que se ha comprobado no tiene ninguna clase de efecto o consecuencia, digamos dulces que tienen la apariencia de pastillas). Después se observa si hubo o no alguna diferencia por lo que respecta a la cura de la enfermedad.

Ahora bien, el hecho de que un grupo no se exponga al tratamiento experimental no significa que su participación en el experimento sea pasiva (que mientras el grupo experimental participa en un cierto tratamiento, el grupo de control puede hacer lo que quiera, v.g., irse a algún otro lado y luego regresar, platicar entre si o seguir una rutina cotidiana). Por el contrario, significa que realiza las mismas actividades que el grupo experimental excepto el someterse al estímulo. Por ejemplo, si el grupo experimental va a ver un programa de televisión con contenido violento, el otro grupo podría ver el mismo programa pero sin las escenas violentas —otras versiones del mismo programa—. Desde luego, en ocasiones resulta muy difícil definir lo que es no exponerse al estímulo. Por ejemplo, si pretendemos probar la efectividad de una nueva psicoterapia, si al grupo de control lo exponemos a una psicoterapia tradicional sería difícil afirmar que su nivel es cero, puesto que es muy probable que ambas psicoterapias tengan algo en común. Por otro lado, si el grupo de control no recibiera ninguna psicoterapia, las diferencias entre los dos grupos bien podrían atribuirse al efecto que puede tener el que las personas participen por primera vez en una psicoterapia y no al efecto de esa nueva psicoterapia.

En general, puede afirmarse en un experimento que, si en ambos grupos todo fue “igual” menos la exposición a la variable independiente, es muy razonable pensar que las diferencias entre los grupos se deban a la presencia-ausencia de la variable independiente.

Más de dos grados

En otras ocasiones, se puede *hacer variar o manipular la variable independiente en cantidades o grados*. Por ejemplo, en el caso del análisis del posible efecto del contenido

antisocial por televisión sobre la conducta agresiva de ciertos niños, podría hacerse que un grupo se expusiera a un programa de televisión sumamente violento (con presencia de violencia física verbal y no verbal —golpes, asesinatos, insultos muy fuertes, etc.—); un segundo grupo se expusiera a un programa medianamente violento (únicamente con violencia verbal —insultos menos fuertes—), y un tercer grupo se expusiera a un programa prosocial. En este ejemplo, se tendrían tres niveles o cantidades de la variable independiente, lo cual puede representarse de la siguiente manera:

X_1	(programa sumamente violento)
X_2	(programa medianamente violento)
	(ausencia de violencia, programa prosocial)

Manipular la variable independiente en varios niveles tiene la ventaja de que no sólo se puede determinar si la presencia de la variable independiente o tratamiento experimental tiene un efecto, sino también se puede determinar si distintos niveles de la variable independiente tienen diferentes efectos. Es decir, si la magnitud del efecto (Y) depende de la intensidad del estímulo (X_1 , X_2 , X_3 , etcétera). Ahora bien, ¿cuántos niveles de variación deben ser incluidos? Una respuesta exacta no puede darse, solamente que debe haber al menos dos niveles de variación y ambos tendrán que diferir entre sí. El problema de investigación, los antecedentes (estudios anteriores) y la experiencia del investigador pueden proveer alguna indicación sobre el número de niveles de variación que necesita ser incorporado en cierto experimento. Y cabría agregar: Entre más niveles mayor información, pero el experimento se va complicando, cada nivel adicional implica un grupo más (Christensen, 1980).

Modalidades de manipulación en lugar de grados

Existe otra forma de manipular una variable independiente que consiste en *exponer a los grupos experimentales a diferentes modalidades de ésta pero sin que ello implique cantidad*. Por ejemplo, supongamos que un investigador quiere probar el efecto que tienen distintas fuentes de retroalimentación sobre el desempeño en la productividad de los trabajadores de una fábrica. La retroalimentación sobre el desempeño se refiere a que le digan a una persona cómo está desempeñándose en su trabajo (qué tan bien o mal lo hace). A un grupo de trabajadores se le proporcionaría retroalimentación sólo mediante su supervisor, a otro grupo la retroalimentación provendría por escrito (sin contacto “cara a cara” con otra persona), y a un tercer grupo se le indicaría que entre los compañeros de trabajo se proporcionarían retroalimentación entre sí (todo en la ejecución de una determinada tarea); y luego se compararía la productividad de los grupos. En este caso no se está manipulando la presencia-ausencia de la variable independiente, ni administrando distintas cantidades de ésta, sino que los grupos se exponen a modalidades de la retroalimentación del desempeño, no a intensidades. La variación es provocada por categorías distintas de la variable independiente que no implican en sí cantidades.

En ocasiones, la manipulación de la variable independiente involucra combinadas cantidades y modalidades de ésta. Por ejemplo, si en el caso anterior tuviéramos un grupo al que se le administrara retroalimentación detallada de su desempeño vía el supervisor (en donde se retroalimentara aún las actividades laborales menos trascendentes), un segundo grupo al que se le administrara retroalimentación dosificada de su desempeño, sólo en las labores más importantes y vía el supervisor, un tercer grupo al que se le administrara retroalimentación detallada de su desempeño vía un medio escrito, un cuarto grupo con retroalimentación dosificada por conducto de un medio escrito y un quinto grupo sin retroalimentación. Esto es, se combinan grado (retroalimentación detallada, retroalimentación dosificada y ausencia de retroalimentación) y modalidad (vía superior y medio escrito).

Finalmente, es necesario insistir que *cada nivel o modalidad implica — al menos— un grupo.* Si se tienen tres niveles (grados) o modalidades, se tendrán tres grupos como mínimo.

6.5. ¿CÓMO SE DEFINE LA MANERA EN QUE SE MANIPULARÁN LAS VARIABLES INDEPENDIENTES?

Al manipular una variable independiente es necesario especificar qué se va a entender por esa variable en nuestro experimento. Es decir, trasladar el concepto teórico en un estímulo experimental, en una serie de operaciones y actividades concretas a realizar. Por ejemplo, si la variable independiente a manipular es la exposición a la violencia televisada, el investigador debe pensar en cómo va a transformar ese concepto en una serie de operaciones experimentales. En este caso podría ser: ““la violencia televisada será operacionalizada (transportada a la realidad) como ver un programa en donde haya riñas y golpes, insultos, agresiones, uso de armas de fuego, crímenes o intentos de crímenes, azotes de puertas, se aterre a personas, persecuciones, etcétera”. Entonces se selecciona un programa donde se muestren tales conductas (v.g., “Miami Vice”, “El Justiciero”, ““Magnum”, “El Precio del Deber” o una telenovela mexicana, brasileña o venezolana donde se presenten dichos comportamientos). El concepto abstracto ha sido convertido en un hecho real. Si la variable independiente es la orientación principal del profesor hacia la autonomía o el control, debemos definir qué comportamientos concretos, filosofía, instrucciones al grupo, presentación, personalidad, etcétera, debe mostrar cada tipo de profesor (y analizar sus diferencias). Si la variable independiente es el tipo de psicoterapia recibida (y se tienen tres tipos, esto es, tres grupos), debemos definir muy específicamente y con lujo de detalles en qué va a consistir cada psicoterapia.

EJEMPLO

Cátedra “Metodología para la investigación en Ciencia Política”. Unidad IV
Hernández Sampieri, R. Collado, L. Lucio, P. Metodología de la investigación (MacGraw Hill, México)

Naves y Poplawsky (1984) diseñaron un experimento para poner a prueba la siguiente hipótesis: “A mayor grado de información sobre la deficiencia mental que el sujeto normal maneje, mostrará menos evitación en la interacción con el deficiente mental”.

La variable independiente era —pues— “el grado de información sobre la deficiencia mental”, y la dependiente “la conducta de evitación en interacciones con deficientes mentales”. La primera fue manipulada mediante dos niveles de información: 1) información cultural y 2) información sociopsicológica. Por lo tanto, había dos grupos: uno con información cultural y otro con información sociopsicológica. El primer grupo no recibió ningún tipo de información sobre la deficiencia mental, ya que se supuso “que todo individuo por pertenecer a cierta cultura, maneja este tipo de información y está conformada por nociones generales y normalmente estereotipadas sobre la deficiencia mental; de ello se desprende que si un sujeto basa sus predicciones sobre la conducta del otro en el nivel cultural, obtendrá mínima precisión y pocas probabilidades de controlar el evento comunicativo” (Naves y Poplawsky, 1984, p. 119).

El segundo grupo acudió a un centro de capacitación para deficientes mentales, en donde tuvo una reunión con dichos deficientes, quienes les proporcionaron información sociopsicológica sobre ellos (algunos deficientes contaron sus problemas en el trabajo y sus relaciones con superiores y compañeros, se trataron temas como el amor y la amistad). Asimismo, se intercambiaron experiencias más personales. Este grupo pudo observar lo que es la deficiencia mental, la manera como se trata clínicamente y los efectos sobre la vida cotidiana de quien la padece. Recibió información sociopsicológica.

Después, todos los sujetos eran expuestos a una interacción sorpresiva con un supuesto deficiente mental (que en realidad era un actor entrenado para ello y con conocimientos sobre la deficiencia mental). La situación experimental estaba bajo riguroso control y se filmaban las interacciones para medir el grado de evitación hacia el deficiente, a través de cuatro dimensiones: a) distancia física, b) movimientos corporales que denotaban tensión, c) conducta visual y d) conducta verbal. Cabe mencionar que se comprobó la hipótesis, el grupo con información cultural mostró mucho mayor conducta de evitación que el grupo con información sociopsicológica.

El punto es que un concepto teórico (grado de información sobre la deficiencia mental) fue traducido en la práctica a dos niveles de manipulación experimental.

Dificultades para definir cómo se manipularán las Variables independientes

En ocasiones no resulta difícil trasladar el concepto teórico (variable independiente) en operaciones prácticas de manipulación (tratamientos o estímulos experimentales). Por ejemplo, si se busca analizar el efecto de utilizar distintas apelaciones publicitarias para promover medicamentos —emotivas versus racionales— sobre la predisposición para comprarlos, la variable independiente podría operacionalizarse de la siguiente manera: se elabora un comercial de televisión sobre un medicamento en particular, en el cual el argumento de venta es que se trata de un producto que ha sido sometido a pruebas científicas de laboratorio y que demostró su efectividad, además de que es recomendado por tal y tal asociación médica (apelaciones racionales); y se elabora otro comercial cuyo argumento de

ventas es que el medicamento es tradición en muchas familias y desde nuestros abuelos se utilizaba (apelación emotiva). Los modelos de ambos tipos de comerciales son los mismos, los dos son a color, duran 30 segundos y en fin, la única diferencia es la apelación, tanto en el nivel verbal como en el no verbal. Un grupo se expone a la manipulación racional y el otro a la emotiva, por lo demás las condiciones de exposición son similares, y después se analiza el impacto de la manipulación en la variable dependiente. Manipular la paga (cantidades de dinero otorgadas), la retroalimentación, el reforzamiento y la administración de un medicamento no es tan difícil.

Sin embargo, a veces es sumamente complicado representar el concepto teórico en la realidad, sobre todo con variables internas, variables que pueden tener varios significados o variables que sean difíciles de alterar. La socialización, la cohesión, la conformidad, el poder, la motivación individual y la agresión son conceptos que requieren de un enorme esfuerzo por parte del investigador para ser operacionalizados.

Guía para sortear las dificultades

Para definir cómo se va a manipular una variable es necesario:

1. *Que se consulten experimentos antecedentes* para ver si en éstos resultó la forma de manipular la variable. Al respecto, resulta imprescindible analizar si las manipulaciones de esos experimentos pueden aplicarse al contexto específico del nuestro o cómo pueden ser extrapoladas a nuestra situación experimental. Por ejemplo, Bylenga (1977), en un experimento para estudiar la relación entre las recompensas extrínsecas y la motivación intrínseca, hizo que los sujetos —como parte de la manipulación— jugaran al “juego del ahorcado” en una terminal de computadora. En este juego hay que adivinar palabras. Unos sujetos recibían dinero por su buen desempeño y otros no, a unos se les felicitaba por sus logros y a otros no. Hernández-Sampieri y Cortés (1982) replicaron el experimento en México, y como no dispusieron de un programa de computadora que incluyera dicho juego, tuvieron que construir una máquina de madera mecánica que realizaba las mismas funciones que el programa de computadora.

2. *Que se evalúe la manipulación antes de que se conduzca el experimento.* Hay varias preguntas que el experimentador debe hacerse para evaluar su manipulación antes de llevarla a cabo: ¿Mis operaciones experimentales representan la variable conceptual que tengo en mente?, ¿los diferentes niveles de variación de la variable independiente harán que los sujetos se comporten diferente? (Christensen, 1980), ¿qué otras maneras hay de manipular a esta variable?, ¿mi manera es la mejor? Si el concepto teórico no es trasladado adecuadamente a la realidad, lo que sucede es que finalmente haremos otro experimento distinto al que pretendemos. Por ejemplo, si deseáramos averiguar el efecto de la ansiedad sobre la memorización de conceptos y nuestra manipulación es errónea (en lugar de provocar ansiedad, genera inconformidad), los resultados del experimento nos ayudarán tal vez a explicar la relación inconformidad-memorización de conceptos, pero de ninguna

manera servirán para analizar el efecto de la ansiedad en dicha memorización. Y a lo mejor no nos damos cuenta y creemos que aportamos algo cuando en realidad no lo hicimos. Asimismo, en muchos casos cuando la manipulación resulta errónea y generamos experimentalmente otra variable independiente distinta a la que nos interesa, ni siquiera se aporta evidencia sobre el efecto de esa otra variable porque no se estaba preparado para ello. En resumen, si la manipulación es errónea nos puede pasar que: 1) el experimento no sirva para nada, 2) vivamos en el error (creer que “algo es” cuando verdaderamente “no lo es”) y 3) tengamos resultados que no nos interesan (si nos interesaran habríamos pensado en otro experimento).

También, si la presencia de la variable independiente en el(los) grupo(s) experimental(es) es débil probablemente no se encontrarán efectos, pero no porque no puedan haberlos. Por ejemplo, si pretendemos manipular la violencia televisada y nuestro programa no es realmente violento (incluye uno que otro insulto y algunas sugerencias de violencia física) y no encontramos un efecto, realmente no podemos afirmar o negar que hay un efecto, porque la manipulación fue débil.

3. *Que se incluyan verificaciones para la manipulación.* Cuando se utilizan seres humanos hay varias formas de verificar si realmente funcionó la manipulación (Christensen, 1980). La primera es entrevistar a los sujetos. Por ejemplo, si a través de la manipulación pretendemos generar que un grupo esté altamente motivado hacia una tarea o actividad y el otro no, después del experimento podemos entrevistar a los participantes para ver si el grupo que debía estar muy motivado realmente lo estuvo y el grupo que no debía estar motivado no lo estuvo. Una segunda forma es incluir mediciones relativas a la manipulación durante el experimento. Por ejemplo, aplicar una escala de motivación a ambos grupos cuando supuestamente unos deben estar muy motivados y otros no.

6.6. ¿CUÁL ES EL SEGUNDO REQUISITO DE UN EXPERIMENTO “PURO”?

El segundo requisito es medir el efecto de la variable independiente sobre la variable dependiente. Esto es igualmente importante y como en ella se observa dicho efecto, es requisito que su medición sea válida y confiable. Porque si no podemos asegurar que estuvo adecuadamente medida, los resultados no servirán y el experimento será una pérdida de tiempo. Imaginemos que conducimos un experimento para evaluar el efecto de un nuevo tipo de enseñanza sobre la comprensión de conceptos políticos en ciertos niños, y en lugar de medir comprensión medimos nada más memorización, por más correcta que resulte la manipulación de la variable independiente, como la medición de la dependiente no es válida, el experimento resulta un fracaso. O supongamos que tenemos dos grupos a comparar con mediciones distintas —como por ejemplo preguntas diferentes—, si encontramos diferencias ya no sabremos si se debieron a la manipulación de la independiente o a que se aplicaron exámenes distintos. En el capítulo sobre “elaboración de los instrumentos de medición” se

comenta qué requisitos se necesitan para medir correctamente a una variable, los cuales se aplican también a la medición de la(s) variable(s) dependiente(s) en un experimento. Asimismo, ésta(s) puede(n) ser medida(s) de diversa(s) manera(s); cuestionarios, escalas, observación, entrevistas, mediciones fisiológicas, análisis de contenido, etcétera; las cuales se explican en el mencionado capítulo.

En la planeación de un experimento se debe precisar cómo se van a manipular las variables independientes y cómo a medir las dependientes.

6.7. ¿CUÁNTAS VARIABLES INDEPENDIENTES Y DEPENDIENTES DEBEN INCLUIRSE EN UN EXPERIMENTO?

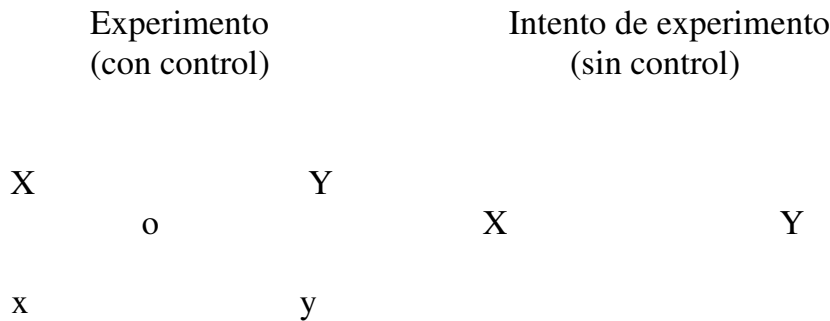
No hay reglas para ello; depende de cómo haya sido planteado el problema de investigación y las limitaciones que se tengan. Por ejemplo, el investigador que estaba interesado en analizar cómo afecta el utilizar apelaciones emotivas versus racionales en comerciales televisivos sobre la predisposición de compra de un producto, solamente se interesa en este problema, entonces tiene una única variable independiente y una sola dependiente. Pero si a él también le interesara analizar el efecto que tenga el utilizar comerciales en blanco y negro versus en color, entonces agregaría esta variable independiente y la manipularía. Tendría dos variables independientes y una dependiente, y cuatro grupos: A) grupo expuesto a apelación emotiva y comercial en blanco y negro, B) grupo expuesto a apelación emotiva y comercial en color, C) grupo expuesto a apelación racional y comercial en blanco y negro, D) grupo expuesto a apelación racional y comercial en color. O podría agregar una tercera independiente: duración de los comerciales, y una cuarta: realidad de los modelos del comercial (personas vivas en contraposición a caricaturas) y así sucesivamente. Claro está que, conforme se aumenta el número de variables independientes, aumentan las manipulaciones que deben hacerse y el número de grupos requeridos para el experimento. Y entraría en juego el segundo factor mencionado (limitantes), a lo mejor no puede conseguir las suficientes personas para tener el número de grupos que requiere o el presupuesto para producir tantos comerciales.

Por otra parte, podría decidir en cada caso (con una, dos, tres o más variables independientes) medir más de una variable dependiente para ver el efecto de la(s) independiente(s) en distintas variables. Por ejemplo, medir, además de la predisposición de compra, la recordación del comercial y la evaluación estética de éste. Otro investigador podría interesarse en evaluar el efecto de una mayor o menor paga en el desempeño de una tarea laboral en diversas variables: motivación intrínseca, productividad, satisfacción laboral y autoestima. Resulta obvio que, al aumentar las variables dependientes, no tienen que aumentarse grupos, porque estas variables no se manipulan. Lo que aumenta es el tamaño de la medición (cuestionarios con más preguntas, mayor número de observaciones, entrevistas

más largas, etcétera) porque hay más variables que medir.

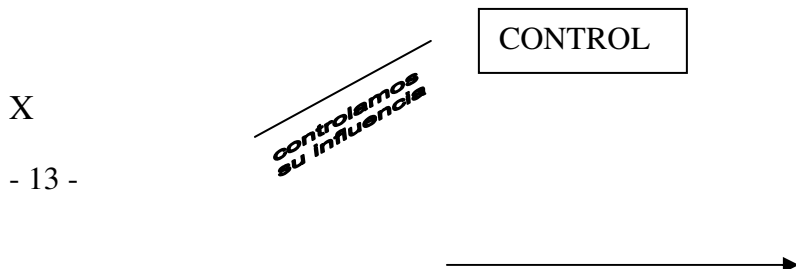
6.8. ¿CUAL ES EL TERCER REQUISITO DE UN EXPERIMENTO “PURO”?

El tercer requisito que todo experimento “verdadero” debe cumplir es el *control o validez interna de la situación experimental*. El término “control” tiene diversas connotaciones dentro de la experimentación; sin embargo, su acepción más común se refiere a que si se observa con el experimento que una o más variables independientes al ser manipuladas hacen variar a la(s) dependiente(s), la variación de estas últimas se deba a la manipulación de la(s) independiente(s) y no a otros factores o causas; o si se observa que una o más independientes no tienen un efecto sobre la(s) dependiente(s), se pueda estar seguro de ello. En términos más coloquiales, *tener ‘control’ significa saber qué está ocurriendo realmente con la relación entre la(s) independiente(s) y la(s) dependiente(s)*. Esto podría ilustrarse de la siguiente manera:



Cuando hay control podemos conocer la relación causal, cuando no se logra el control no se puede conocer dicha relación (no se sabe qué está detrás del “cuadro negro”, podría ser por ejemplo: X—Y, o X Y —no hay relación—). En la estrategia de la investigación experimental, “el investigador no manipula una variable sólo para comprobar lo que le ocurre con la otra, sino que, al efectuar un experimento, es necesario realizar una observación controlada” (Van Dalen y Meyer, 1984).

Dicho de una tercer manera, *lograr ‘control’ en un experimento es controlar la influencia de otras variables extrañas —que no son de nuestro interés— sobre la(s) variable(s) dependiente(s), para que así podamos saber realmente si la(s) variable(s) independiente(s) que nos interesa(n) tiene(n) o no efecto en la(s) dependiente(s)*. Ello podría esquematizarse así:



X	(extrañas)		
X	(de interés, variable independiente manipulada)	vemos su efecto o ausencia de éste	Y (variable dependiente medida)

Es decir, “purificamos” la relación de X (independiente) con Y (dependiente) de otras posibles fuentes que afecten a Y y que nos “contaminan” el experimento. Aislamos la(s) relación (es) que nos interesa (n). Por ejemplo, si deseamos analizar —a través de un experimento— el efecto que pueda tener un comercial sobre la predisposición de compra hacia el producto que anuncia el comercial, sabemos que puede haber otras razones o causas por las que las personas piensen en comprar el producto (la calidad del producto, su precio, cualidades, prestigio de la marca, etcétera). Entonces, el experimento deberá controlar la posible influencia de estas otras causas para que así se pueda saber si el comercial tiene o no un efecto. De otra manera, si se observa que la predisposición de compra es elevada y no hay control, no sabremos si el comercial es la causa o lo son las demás causas posibles.

Lo mismo ocurre con un método de enseñanza, cuando por medio de un experimento se desea evaluar su influencia en el aprendizaje. Si no hay control, no sabremos si un elevado aprendizaje se debió al método o a que los sujetos eran sumamente inteligentes, o bien que éstos tenían conocimientos aceptables de los contenidos o a cualquier otro motivo. También, si no hay aprendizaje no sabremos si esto se debe a que los sujetos estaban muy desmotivados hacia los contenidos a enseñar, que eran poco inteligentes o quién sabe a qué otra causa se pudo deber.

De hecho, el control lo buscamos aún para explicar muchas situaciones cotidianas. Cuando —por ejemplo— alguien nos atrae y tiene un gesto agradable con uno en un momento dado (v.g., tratarnos muy cortésmente), pensamos en la posibilidad de que uno también le sea atractivo a esa persona y buscamos obtener “control” sobre nuestra explicación de por qué esa persona es tan cortés con uno. Es decir, buscamos descartar otras posibles explicaciones para evaluar si nuestra explicación es o no la correcta. Tratamos de descartar explicaciones tales como: “lo que sucede es que esa persona en sí es muy cortés, muy educada”; “lo que ocurre es que esa persona quiere obtener algo de mí”, “tal vez le recordé a alguien que le simpatizaba”, “fue casual su comportamiento”, “estaba de muy buen humor”, etcétera.

Fuentes de invalidación interna

Existen diversos factores o fuentes que pueden hacer que nos confundamos y ya no sepamos si la presencia de una variable independiente o tratamiento experimental tiene o no

un verdadero efecto. Se trata de explicaciones rivales a la explicación de que la variable independiente o independientes afecta(n) a la(s) dependiente(s). En un libro clásico de Campbell y Stanley (1966), se definieron estas explicaciones rivales, las cuales han sido ampliadas y se han agregado otras en referencias más recientes (v.g., Campbell —1975—, Matheson, Bruce y Beauchamp —1985—; Christensen —1980—, Babbie —1979—). A estas explicaciones rivales se les conoce como *fuentes de invalidación interna porque atentan contra la validez interna de un experimento. Ésta se refiere a qué tanta confianza tenemos en que los resultados del experimento sea posible interpretarlos y éstos sean válidos. La validez interna tiene que ver 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), cuando las mediciones de la variable dependiente son confiables y válidas, y cuando el análisis es el adecuado de acuerdo al 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) y Babbie (1979).

1. *Historia.* Son acontecimientos que ocurren durante el desarrollo del experimento que afectan a la variable dependiente y pueden confundir los resultados experimentales. Por ejemplo, si durante un experimento para analizar el efecto que provocan distintas formas de retroalimentación en el trabajo sobre la motivación, teniendo dos grupos de obreros, a uno de éstos le aumentan el salario o se le reúne y felicita por su desempeño en el trabajo mientras está llevándose a cabo el experimento. 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 que afectan los resultados del experimento, como el cansancio, hambre, aburrición, aumento en la edad y cuestiones similares. Por ejemplo, si a un grupo de niños se le expone a un nuevo método de aprendizaje durante dos años, los resultados pueden estar influidos simplemente por la maduración de los niños durante el tiempo que duró el experimento. En un experimento los sujetos pueden cansarse y sus respuestas pueden estar afectadas por ello.
3. *Inestabilidad.* Poca o nula confiabilidad de las mediciones, fluctuaciones en las personas seleccionadas o componentes del experimento, o inestabilidad autónoma de mediciones repetidas aparentemente “equivalentes”. Imaginemos que en un experimento sobre memorización se tienen dos grupos y a uno —al sortearlos— le toca en la mañana y a otro en la tarde, y el experimento requiriera de elevada concentración por parte de los sujetos que participan en él, pudiera ser posible que la concentración del grupo de la mañana fuera diferente del de la tarde (en la mañana habría más ruido, transitar de personas, etcétera); y ello podría afectar la interpretación de los resultados. A lo mejor, diferencias en los grupos se deben a diferencias 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.

4. *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 de clases socioeconómicas en el cual tenemos un grupo, a éste se les aplica una prueba de prejuicio para ver su nivel de prejuicio, luego se les expone a un estímulo experimental —que supuestamente debe disminuir el nivel de prejuicio—, y posteriormente se vuelve a medir el prejuicio para evaluar si disminuyó o no. Puede ocurrir que las calificaciones o puntuaciones de prejuicio de la segunda medición (después del estímulo) se encuentren influidas por la aplicación de la primera prueba de prejuicio (antes del estímulo). Es decir, la administración de la primera prueba puede sensibilizar a los participantes del experimento y cuando respondan a la segunda prueba, sus respuestas podrían estar afectadas por esa sensibilización. Si disminuye el prejuicio y no hay control, no sabremos cuánto se debió al estímulo experimental o variable independiente y cuánto a dicha sensibilización. Cuántas veces no nos ha pasado que cuando nos aplican una prueba decimos: ¡Esta prueba ya la he contestado varias veces!

5. *Instrumentación.* Esta fuente hace referencia a cambios en los instrumentos de medición o en los observadores participantes que pueden producir variaciones en los resultados que se obtengan. 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 efectividad de cada método y comparar. Si los exámenes no fueran equivalentes podría presentarse la instrumentación. Imaginemos que uno es más fácil que el otro, ¿cómo podremos estar 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 distintos exámenes?

6. *Regresión estadística.* Esta fuente se refiere a un efecto provocado por una tendencia de sujetos seleccionados sobre la base de puntuaciones extremas, a regresar —en pruebas posteriores— a un promedio en la variable en la que fueron seleccionados. Por ejemplo, si pretendemos evaluar el efecto del liderazgo autocrático del profesor sobre 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 los sujetos se encontrarán sumamente ansiosos durante la aplicación de la primer prueba (porque tienen un examen difícil al día siguiente). Podría ocurrir que se encontrara que están “menos ansiosos” después de recibir la manipulación, “es decir, aparentemente no hubo efecto”. 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 moverse —esto es, regresar— hacia el promedio de la distribución como función de mediciones repetidas (Christensen, 1981). Entre una primera y una segunda medición, las puntuaciones más altas tienden a bajar y las más bajas a aumentar. Este fenómeno de regresión se puede presentar porque ambas mediciones no están perfectamente correlacionadas. Para un análisis más detallado de este fenómeno se recomienda Campbell y Stanley (1973, Pp. 24-28).

7. *Selección.* Ésta puede presentarse como resultado de elegir a los sujetos para los grupos del experimento, de tal manera que los grupos no sean equiparables. Es decir, si no se seleccionan los sujetos para los grupos asegurándose la equivalencia de éstos, la selección puede resultar tendenciosa. Por ejemplo, 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.

8. *Mortalidad experimental.* Esta fuente se refiere a diferencias en la pérdida de participantes entre los grupos que se comparan. Si en un grupo se pierde —digamos— al 25% de los participantes y en otro grupo sólo al 2%, los resultados pueden verse influidos por ello, además del tratamiento experimental. 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 les expone a un programa prosocial que nadie ha visto. Condiciones agresivas, dolorosas, cansadas, etc.; pueden provocar mortalidad diferencial en los grupos, y ésta puede ocurrir no sólo por el experimento en sí sino el tipo de personas que componen cada grupo o factores externos al experimento.

9. *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. La selección resulta en diferentes tasas de maduración o cambio autónomo entre grupos. Por ejemplo, si seleccionamos para un grupo sujetos que acostumbran alimentarse a cierta hora (12:00 p.m.) y para el otro personas que se alimentan a otra hora (3:00p.m.), y el experimento se lleva a cabo de 11:00 a.m. a 14:30 p.m., la selección tendenciosa provocará un efecto de maduración distinto en los dos grupos: hambre. Esto podría ocurrir si en una convención de una empresa multinacional, experimentamos con ejecutivos de distintos países latinoamericanos y no tomamos en cuenta la interacción que pueda darse entre la selección y la maduración.

10. *Otras interacciones.* Asimismo, pueden darse diversos efectos provocados por la interacción de las fuentes de invalidación interna mencionadas. La selección puede interactuar con la mortalidad experimental, la historia con la maduración, la maduración con la inestabilidad, etcétera. También pueden afectar varias de estas fuentes y la validez interna se deteriora en mayor grado. Cada vez estamos más inseguros respecto a la(s) causa(s) que motivó cambios en la(s) variable(s) dependiente(s); o si no hubo cambios, no sabemos si se debió a que una o más fuentes contrarrestaron los posibles efectos. Por ejemplo, el método de enseñanza más exitoso se utilizó con los menos motivados, y el menos exitoso con los más motivados, y ello provoca que se compensen ambos factores. No sabemos cómo interpretar los resultados.

Por todo lo anterior es necesario eliminar a estas fuentes de invalidación interna mediante el control, para así poder conocer el efecto de la variable independiente (o las independientes) sobre la dependiente (o dependientes).

El experimentador como fuente de invalidación interna

Otra posible razón que puede atentar contra la interpretación correcta y certera de los resultados de un experimento *es la interacción entre los sujetos y el experimentador*, la cual puede ocurrir de diferentes formas. Los sujetos pueden entrar al experimento con ciertas actitudes, expectativas y prejuicios que pueden alterar su comportamiento durante el estudio. Por ejemplo, no colaborar y ser críticos negativamente, hasta el punto de llegar a ser hostiles. Ello debe tenerse en mente antes y durante la investigación. Debe analizarse qué sujetos pueden arruinar el experimento y eliminarse o en todos los grupos debe haber personas con actitudes positivas y negativas (si quienes tienen actitudes negativas van a un único grupo, la validez interna estará en problemas). Recordemos que las personas que participan 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 puede afectar los resultados de la investigación*. El experimentador también no es un observador pasivo que no interactúa, sino un observador activo que puede influir los resultados del estudio (Christensen, 1981). El experimentador tiene una serie de motivos que lo llevan a realizar su experimento y desea probar que lo que hipotetiza se demuestre a través de éste. Desea comprobar su 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. Lo anterior debe evitarse, y en vados casos quien debe tratar con los sujetos no debe ser el experimentador sino alguien que no conozca la hipótesis, las condiciones experimentales ni los propósitos del estudio, simplemente que se le den instrucciones precisas sobre lo que debe hacer y cómo hacerlo.

Tampoco los sujetos que participan en el experimento deben conocer las hipótesis y condiciones experimentales e incluso frecuentemente es necesario distraerlos de los verdaderos propósitos del experimento (aunque, desde luego, al finalizar éste se les debe dar una explicación completa del experimento). Cuando —por ejemplo— se analizan efectos de medicamentos, los investigadores hacen creer a un grupo que se les está administrando medicamentos cuando en realidad no es así, se les dan píldoras de azúcar. Esto evita la influencia que pudiera tener la expectativa de recibir medicamentos sobre la variable dependiente. A esta sustancia que no tiene efectos se le suele denominar “placebo”. Con métodos de instrucción —por ejemplo— puede ocurrir que el grupo que se habrá de exponer al método innovador sea 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, etcétera. Esto debe de tomarse en cuenta.

6.9. ¿CÓMO SE LOGRA EL CONTROL Y LA VALIDEZ INTERNA?

El control en un experimento logra la validez interna, y el control se alcanza

mediante: 1) *varios grupos de comparación* (dos como mínimo) y 2) *equivalencia de los grupos en todo, excepto la manipulación de la variable o las variables independientes.*

6.9.1. Varios grupos de comparación

Es necesario que en un experimento se tengan por lo menos dos grupos que comparar. En primer término, porque si nada más se tiene un grupo, no se puede saber si influyeron las fuentes de invalidación interna o no. Por ejemplo, si mediante un experimento buscamos probar la hipótesis de que: “A mayor información psicológica sobre una clase social, menor prejuicio hacia esta clase”; y decidimos tener un solo grupo en el experimento. Este grupo se expondría a un programa televisivo donde se proporcione información sobre la manera como vive dicha clase, sus angustias y problemas, necesidades, sentimientos, aportaciones a la sociedad, etcétera; para luego observar el nivel de prejuicio. Este experimento podría esquematizarse así:

<i>Momento 1</i>	<i>Momento 2</i>
Exposición al programa televisivo	Observación del nivel de prejuicio

Todo en un único grupo. ¿Qué sucede si se observa un bajo nivel de prejuicio en el grupo?, ¿podemos deducir que se debió al estímulo? Desde luego que no. Es posible que —efectivamente— el nivel bajo de prejuicio se deba al programa de televisión, que es la forma de manipular la variable independiente “información psicológica sobre una clase social”, pero también pudiera deberse a que los sujetos tenían un bajo nivel de prejuicio antes del experimento y —en realidad— el programa no afecta. Y no lo podemos saber porque no hay una medición del nivel de prejuicio al inicio del experimento (antes de la presentación del estímulo experimental), es decir, no existe punto de comparación. Y aunque hubiera ese punto de comparación inicial, con un solo grupo no podremos estar seguros de cuál fue la causa del elevado prejuicio. Supongamos que el nivel de prejuicio antes del estímulo era alto y después del estímulo bajo. Pudiera ser que el estímulo es la causa del cambio, pero también pudo ocurrir lo siguiente:

1. Que la primera prueba de prejuicio sensibilizara a los participantes y sus respuestas a la segunda prueba fueran influidas por aquélla. Así, las personas se concientizaron de lo negativo de ser prejuiciosas como resultado de responder a la primer prueba.
2. Que los sujetos seleccionados se cansaron durante el experimento y sus respuestas a la segunda prueba fueron “a la ligera” (maduración).
3. Que había ocurrido un suceso antes del experimento que los prejuicio —momentáneamente— hacia esa clase social (una violación en la localidad a cargo de un individuo de esa clase), pero después “regresaron” a su nivel de prejuicio normal.
4. Que durante el experimento se salieron los sujetos prejuiciosos o parte importante de

ellos.

O bien otras razones. Y si no se hubiera encontrado un cambio en el nivel de prejuicio entre la primera prueba (antes del programa) y la segunda (después del programa), esto podría significar que la exposición al programa no tienen efectos, pero también podría ser que el grupo seleccionado es muy prejuicioso y a lo mejor el programa sí tiene efectos en personas con niveles comunes de prejuicio. Incluso, podría haber otras explicaciones alternativas. Con un solo grupo no podemos estar seguros de que los resultados se deben al estímulo experimental o a otras razones, siempre quedará la duda, los “experimentos” con un grupo se basan en sospechas o en lo que “aparentemente es”, pero faltan fundamentos. Al tener un único grupo se corre el riesgo de seleccionar sujetos atípicos (los más inteligentes al experimentar con métodos de enseñanza, los trabajadores más motivados al experimentar con programas motivacionales, los consumidores más acríticos, las parejas de novios más integradas, etcétera) y el riesgo de que actúen la historia, maduración, administración de prueba, instrumentación y demás fuentes de invalidación interna; sin que el experimentador se dé cuenta.

Por ello, *el investigador debe tener como mínimo un punto de comparación: dos grupos, uno al que se le administra el estímulo y otro al que no, a este último recordemos que se le denomina grupo de control,*²² y tal como se mencionó al hablar de manipulación, *a veces se requiere tener de varios grupos, cuando se desea averiguar el efecto de distintos niveles de la variable independiente.*

6.9.2. Equivalencia de los grupos

Pero para tener control no basta con tener dos o más grupos, sino que éstos deben ser similares en todo, excepto la manipulación de la variable independiente. El control implica que todo permanece constante menos esta manipulación. Si entre los grupos que conforman el experimento todo es similar o equivalente, excepto la manipulación de la independiente, diferencias entre los grupos pueden atribuirse a ella y no a otros factores (entre los cuales están las fuentes de invalidación interna). Por ejemplo, si tenemos dos grupos de sustancias ““A”, ““B”, ““C” y ““D” para mezclarlas en dos recipientes ““1” y ““2”. La sustancia ““A” es la misma para cada recipiente (y en igual cantidad), lo mismo que las otras tres sustancias. Los recipientes tienen el mismo volumen y son idénticos. La presión y temperatura en ambos es igual. Los instrumentos para mezclar son también iguales. El lugar y la atmósfera son equivalentes. Y en fin, todo permanece constante. El resultado final: la mezcla, tendrá que ser la misma (idénticas características). Pero, si algo se hace variar o se manipula, si es distinto al hacer ambas mezclas en los recipientes “1” y “2” (digamos que a una de las mezclas se le agrega la sustancia ““E” y a la otra no).

Las diferencias entre las mezclas pueden atribuirse a la presencia de esta nueva sustancia, porque todo es igual con excepción de que una mezcla contiene la sustancia ““E”

mientras que la otra no. También puede decirse que, si hay cambios en la mezcla con la nueva sustancia, éstos pueden atribuirse a la sustancia “E”. Sin embargo, para poder llegar a esta conclusión fue necesario un grupo de comparación equivalente. Si nada más se tuviera una mezcla, no podríamos estar seguros que fue la adición de la sustancia la que provocó cambios, pudo haber sido que la reacción de los primeros cuatro elementos fue tardía y el quinto nada tuvo que ver, o que un aumento en la temperatura fue lo que provocó el cambio u otras cuestiones. Desde luego, en las ciencias químicas se hace una mezcla y se observa lo que ocurre, se hace otra mezcla y se observa, se hace una tercera mezcla a la que se le agrega otra sustancia y se continúa observando, y así sucesivamente. Y si no encontramos diferencias entre la mezcla con la sustancia “E” y la que no la tiene, decimos que “E” no tiene efecto.

Lo mismo debe ser en la experimentación de la conducta humana, debemos tener varios grupos de comparación. Imaginemos que queremos probar si una serie de programas de televisión educativos para niños genera mayor aprendizaje en comparación con un método educativo tradicional. Un grupo recibe la enseñanza por medio de los programas, otro grupo

²² El grupo de control es útil precisamente para tener un punto de comparación. Sin él, no podremos saber qué sucede cuando la variable independiente está ausente. Su nombre indica su función: ayudar a establecer el control, colaborando en la eliminación de hipótesis rivales o influencias de las posibles fuentes de invalidación interna.

la recibe por medio de instrucción oral tradicional y un tercer grupo dedica ese mismo tiempo a jugar libremente en el salón de clases. Supongamos que los niños que aprendieron mediante los programas obtienen las mejores calificaciones en una prueba de conocimientos relativa a los contenidos enseñados, los que recibieron el método tradicional obtienen calificaciones mucho más bajas y los que jugaron obtienen puntuaciones de cero o cerca de este valor. Aparentemente, los programas son mejor medio de enseñanza que la instrucción oral. Pero si los grupos no son equivalentes, por ejemplo: los niños más inteligentes, estudiosos y con mayor empeño fueron asignados al grupo que fue instruido por televisión, o simplemente su promedio de inteligencia y aprovechamiento era mas elevado; o la instructora del método tradicional no era una buena maestra, o los niños expuestos a este método tenían mayor carga de trabajo y exámenes los días en que se desarrolló el experimento, etcétera; no podemos confiar en que las diferencias se deben realmente a la manipulación de la variable independiente y no a otros factores, o a la combinación de ambos. ¿Cuánto se debió al método y cuánto a otros factores? Para el investigador la respuesta a esta pregunta se convierte en un enigma. No hay control.

Si experimentamos con métodos de motivación para trabajadores, y a un grupo enviáramos a los que laboran en el turno matutino, mientras que al otro grupo enviáramos a los del turno vespertino; ¿quién nos asegura que antes de iniciar el experimento ambos tipos de trabajadores están igualmente motivados? Pudo haber diferencias en la motivación inicial porque los supervisores motivan de diferente manera y grado o tal vez los del turno vespertino preferirían trabajar en la mañana o se les pagan menos horas extra, etcétera. Si no

están igualmente motivados podría ocurrir que el estímulo aplicado a los del turno de la mañana aparentara ser el más efectivo, cuando en realidad no es así. Los grupos deben ser equivalentes en todo, excepto en la manipulación de la variable independiente.

Veamos un ejemplo que, si bien es extremo, nos ilustra el impacto tan negativo que puede tener la no equivalencia de los grupos sobre los resultados de un experimento. ¿Qué investigador probaría el efecto de diferentes métodos para disminuir el prejuicio teniendo por una parte a un grupo de miembros del Ku-Klux-Klan, por otra parte a un grupo de seguidores del ideólogo Martin Luther King y un tercer grupo de fascistas. Constituyendo cada grupo, un grupo del experimento.

Los grupos deben ser: Inicialmente equivalentes y equivalentes durante todo el desarrollo del experimento, menos por lo que respecta a la variable independiente. Asimismo, los instrumentos de medición tienen que ser iguales y aplicados de la misma manera.

Equivalencia inicial

Implica que *los grupos son similares entre si al momento de iniciarse el experimento*. Por ejemplo, si el experimento es sobre métodos educativos, los grupos deben ser equiparables en cuanto a número de personas, inteligencia, aprovechamiento, disciplina, memoria, sexo, edad, nivel socioeconómico, motivación, alimentación, conocimientos previos, estado de salud física y mental, interés por los contenidos, extroversión, etcétera. Si inicialmente no son equiparables —digamos— en cuanto a motivación o conocimientos previos, diferencias entre los grupos no podrán ser atribuidas con certeza a la manipulación de la variable independiente. Queda la duda si las diferencias se deben a dicha manipulación o a que los grupos no eran inicialmente equivalentes.

La equivalencia inicial no se refiere a equivalencia entre individuos, porque las personas tenemos —por naturaleza— diferencias individuales; *sino a la equivalencia entre grupos* (conjuntos de individuos). Si tenemos dos grupos en un experimento, es indudable que habrá —por ejemplo— personas muy inteligentes en un grupo, pero también debe haberlas en el otro grupo. Si en un grupo hay mujeres, en el otro debe haberlas en la misma proporción. Y así con todas las variables que puedan afectar a la variable dependiente o dependientes —además de la variable independiente—. El promedio de inteligencia, motivación, conocimientos previos, interés por los contenidos, etcétera, debe ser el mismo en los dos grupos. Si bien no exactamente el mismo, no debe haber una diferencia significativa en esas variables entre los grupos.

FIGURA 6.1
ESQUEMA PARA ILUSTRAR LA EQUIVALENCIA INICIAL DE LOS GRUPOS
Al inicio del experimento

Grupo 1

Grupo 2

Cátedra “Metodología para la investigación en Ciencia Política”. Unidad IV
 Hernández Sampieri, R. Collado, L. Lucio, P. Metodología de la investigación (MacGraw Hill, México)

$V_1 = 8$
 $V_2 = 4$
 $V_3 = 6$
 $V_4 = 7.2$
 $V_5 = 10$
 20 mujeres, 21 hombres
 Promedio de edad = 25 años,
 6 meses
 $V_6 = 2$
 $V_k = K$

$V_1 = 7.9$
 $V_2 = 4.1$
 $V_3 = 6$
 $V_4 = 7.4$
 $V_5 = 9.9$
 19 mujeres, 22 hombres
 Promedio de edad = 25 años,
 4 meses
 $V_6 = 2.1$
 $V_k = K$

V = una variable (V_1 = variable 1, V_2 = variable 2...)

Supongamos que todas las variables pudieran medirse de 1 a 10 (es sólo una suposición con fines explicativos); la equivalencia entre grupos podría conceptualizarse como en la figura 6.1.

Veamos ejemplos de equivalencia entre grupos respecto a algunos rasgos físicos — para ilustrar el concepto—.

<i>Grupo 1</i>	<i>Grupo 2</i>
10 hombres de tez morena y ojos cafés.	11 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.	<u>Equivalencia</u>
5 hombres de tez clara y ojos verdes.	12 hombres de tez clara y ojos cafés.
	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 taño 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 figura 6.2 de la página siguiente.

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

Equivalencia durante el experimento

Además, durante el experimento los grupos deben de mantenerse similares en los aspectos que rodean al tratamiento experimental excepto — como ya se ha mencionado— en la manipulación de la variable independiente: mismas instrucciones (salvo variaciones parte de esa manipulación), personas con las que tratan los sujetos, maneras de recibirlos, lugares con características semejantes (iguales objetos en las habitaciones o cuartos, clima, ventilación, sonido ambiental, etcétera), misma duración del experimento, mismo momento, y en fin todo lo que sea parte del experimento. Entre mayor sea la equivalencia durante el desarrollo de éste, mayor control y posibilidad de que si encontramos o no efectos podamos estar seguros de que verdaderamente los hubo o no.

FIGURA 6.2
 EJEMPLO DE DOS GRUPOS NO EQUIVALENTES

<i>Grupo 1</i>		<i>Grupo 2</i>
3 venezolanos 6 colombianos 5 mexicanos 6 norteamericanos 4 ingleses 7 bolivianos 3 italianos 5 israelitas 4 afganos 3 cubanos		1 venezolano 3 brasileñas 2 mexicanos 1 norteamericano 28 franceses 10 ingleses 4 rusos 2 alemanes 5 suizos 2 nicaragüenses 4 egipcios

Cuando trabajamos simultáneamente con varios grupos, es difícil que las personas que dan las instrucciones y vigilan el desarrollo de los grupos sean las mismas, entonces debe buscarse que su tono de voz, apariencia, edad, sexo u otras características que consideremos que puedan afectar los resultados sean iguales o similares, y mediante entrenamiento debe estandarizarse su proceder. También, a veces se dispone de menos cuartos o lugares que de grupos (v.g., se tienen cuatro grupos y sólo dos cuartos). Entonces, la asignación de los grupos a los cuartos y horarios se hace al azar y se procura que los grupos no estén muy espaciados (lo menos distantes que sea posible). En otras ocasiones, los sujetos reciben los tratamientos individualmente, no puede ser simultánea la exposición a éstos. Pero se deben sortear los individuos de tal manera que en un día —en la mañana—

Cátedra “Metodología para la investigación en Ciencia Política”. Unidad IV
 Hernández Sampieri, R. Collado, L. Lucio, P. Metodología de la investigación (MacGraw Hill, México)

personas de todos los grupos participen en el experimento, lo mismo en la tarde y durante el tiempo que sea necesario (los días que dure el experimento). Esto podría esquematizarse así, teniendo tres grupos:

HORA	DIA 1	DIA 2
9:00	S1	S2
	S2	S1
	S3	S3
10:00	S1	S3
	S2	S1
	S3	S2
11:00	S1	S1
	S3	S2
	S2	S3
12:00	S3	S2
	S2	S1
	S1	S3
13:00	S2	S1
	S3	S2
	S1	S3
14:00	S2	S3
	S3	S2
	S1	S1
15:00	S3	S2
	S1	S1
	S2	S3
16:00	S3	S1
	S2	S2
	S1	S3
17:00		

¿Cómo se logra la equivalencia inicial?: Asignación al azar

Existe un método ampliamente difundido para alcanzar dicha equivalencia que se conoce como *asignación aleatoria o al azar de los sujetos a los grupos del experimento* (en inglés, el término equivalente es “*randomization*”). La asignación al azar nos asegura probabilísticamente que dos o más grupos son equivalentes entre sí. Es una técnica de control que tiene como propósito proveer al investigador la seguridad de que variables extrañas, conocidas o desconocidas, no afectarán sistemáticamente los resultados del estudio (Christensen, 1981). Esta técnica debida a Sir Ronald A. Fisher —en los años cuarenta— ha demostrado durante años y pruebas que funciona para hacer equivalentes a grupos. Como

mencionan Cochran y Cox (1980, p. 24): “La aleatorización es en cierta forma análoga a un seguro, por el hecho de que es una precaución contra interferencias que pueden o no ocurrir, y ser o no importantes si ocurren. Generalmente, es aconsejable tomarse el trabajo de aleatorizar, aun cuando no se espere que haya un sesgo importante al dejar de hacerlo”.

La asignación al azar puede llevarse a cabo mediante pedacitos de papel Se escribe el nombre de cada sujeto (o algún tipo de clave que lo identifique) en un pedacito de papel, luego se juntan los pedacitos en algún recipiente, se revuelven y —sin ver— se van sacando para formar los grupos. Por ejemplo, si se tienen dos grupos, las personas con papelitos nones pueden ir al primer grupo y las personas con pares al segundo grupo; o bien, si se tuvieran 80 personas, los primeros 40 papelitos que se saquen van a un grupo y los restantes 40 al otro.

También, cuando se tienen *dos grupos, la aleatorización puede llevarse a cabo utilizando una moneda no cargada*. Se lista a los sujetos y se designa qué lado de la moneda va a significar el grupo 1 y qué lado el grupo 2 (por ejemplo, “cara” = grupo 1 y cruz” = grupo 2). Con cada sujeto se lanza la moneda y dependiendo de si resulta cara o cruz se le asigna a uno u otro grupo. Este procedimiento está limitado a sólo dos grupos, porque las monedas tienen dos caras. Aunque podrían utilizarse dados o cubos, por ejemplo.

Una tercera forma de asignar a los sujetos a los grupos es mediante el *uso de tablas de números aleatorios*. Una tabla de números aleatorios incluye números del 0 al 9, y su secuencia es totalmente al azar (no hay orden, ni patrón o secuencia), la tabla fue generada mediante un programa de computadora. En el apéndice número 5, se muestra una de estas tablas. Primero, se selecciona al azar una página de la tabla (por ejemplo, preguntándole a alguien que diga un número del 1 al X —dependiendo del número de páginas que contenga la tabla o sorteando números—). En la página seleccionada se elige un punto cualquiera (bien numerando columnas o renglones y eligiendo al azar una columna o renglón, o bien cerrando los ojos y colocando la punta de un lápiz sobre algún punto en la página). Posteriormente, se lee una secuencia de dígitos en cualquier dirección (vertical, horizontal o diagonalmente). Una vez que se obtuvo dicha secuencia, se enumeran los nombres de los sujetos por orden alfabético o de acuerdo a un ordenamiento al azar, colocando cada nombre junto a un dígito. Y se pueden asignar los sujetos nones a un grupo y los pares al otro. Lo mismo da asignar los números del 0 al 5 al primer grupo, y los del 6 al 9 al otro grupo. Si tuviéramos cinco grupos, podríamos hacer que los sujetos con 0 y 1 fueran al primer grupo, con 2 y 3 al segundo, 4 y 5 al tercero, 6 y 7 al cuarto, y 8 y 9 al quinto grupo.

La asignación al azar produce control, en virtud de que las variables que requieren ser controladas (variables extrañas y fuentes de invalidación interna) son distribuidas —aproximadamente— de la misma manera en los grupos del experimento. Y puesto que la distribución es bastante igual en todos los grupos, la influencia de otras variables que no sean la independiente, se mantiene constante porque éstas no pueden ejercer ninguna influencia diferencial en la variable dependiente o variables dependientes (Christensen,

1981).

*La asignación aleatoria funciona más entre mayor sea el número de sujetos con que se cuenta para el experimento, es decir, entre mayor sea el tamaño de los grupos. Los autores recomiendan que para cada grupo se tengan —por lo menos— 15 personas.*²³

Un ejercicio para demostrar las bondades de la asignación al azar

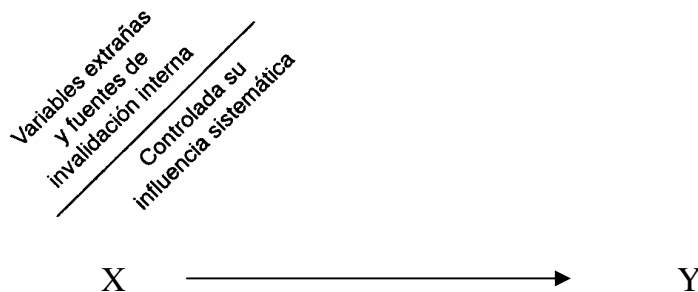
A los estudiantes que se inician en la investigación, a veces les cuesta trabajo creer que la asignación al azar funciona. Para autodemostrarse que sí funciona, es conveniente el siguiente ejercicio:

1. Tómese un grupo de 400 más personas (el salón de clases, un grupo grande de conocidos, etcétera), o imagínese que existe este grupo.
2. Invéntese un experimento que requiera de dos grupos.
3. Imagínese un conjunto de variables que puedan afectar a la(s) variable(s) dependiente(s).
4. Distribuya a cada quien un trozo de papel y pídale que escriban los niveles que tienen en las variables del punto anterior (por ejemplo: sexo, edad, inteligencia, escuela de procedencia, interés por algún deporte, motivación hacia algo —de uno a 10—, etcétera). Las variables pueden ser cualquiera, dependiendo de su ejemplo.
5. Asigne al azar los papelitos a dos grupos.
6. Compare número de mujeres y -hombres en los dos grupos; promedios de inteligencia, edad, motivación, ingreso de su familia o lo que haya pedido. Verá que ambos Verá que ambos grupos son sumamente parecidos”.

²³ Este criterio está basado en la experiencia de los autores.

Si no cuenta con un grupo real, hágalo teóricamente. Usted mismo escriba los valores de las variables en los papelitos y verá cómo los grupos son bastante parecidos (equiparables). Desde luego, normalmente no son “perfectamente iguales” pero sí comparables.

Los resultados de la asignación aleatoria podrían esquematizarse de la siguiente manera:



(variable independiente)

(variable dependiente)

Purificamos la relación. Al controlar todo lo que puede afectar a la variable dependiente (manteniéndolo constante), y hacer variar a la independiente, se puede deducir que los efectos se deben a ésta, o si no hay efectos se puede atribuir que la variable independiente no los tiene.

Si la *única* diferencia que distingue a los grupos experimental y de control es la variable independiente, diferencias entre los grupos pueden atribuirse a ésta. Pero si hay otras diferencias no podemos hacer tal atribución.

Otra técnica para lograr la equivalencia inicial: el emparejamiento

El segundo método para intentar hacer inicialmente equivalentes a los grupos es el emparejamiento o técnica de apareo (en inglés “matching”). Existen diversas modalidades de este método, sin embargo, el más común es el que a continuación se va a describir. El proceso consiste en igualar a los grupos en relación con alguna variable específica, que se piensa puede influir en forma decisiva a la variable dependiente o variables dependientes.

El primer paso es elegir a esa variable de acuerdo con algún criterio teórico. Es obvio que la variable seleccionada debe estar muy relacionada con la(s) variable(s) dependiente(s). Por ejemplo, si se pretendiera analizar el efecto de utilizar distintos tipos de materiales suplementarios de instrucción sobre el desempeño en la lectura, el apareamiento podría hacerse sobre la base de la variable “agudeza visual”. Si el experimento tuviera que ver con el impacto que tienen distintas técnicas de discusión en grupo sobre la efectividad de éste, el emparejamiento podría hacerse tomando en cuenta la “motivación para trabajar en grupo”. Experimentos sobre métodos de enseñanza pueden emparejar a los grupos en “conocimientos previos, aprovechamiento anterior en una asignatura relacionada con los contenidos a enseñar” o “inteligencia”. Experimentos relacionados con actitudes hacia productos o conducta de compra pueden utilizar para aparear a los grupos, la variable “ingreso”. En cada caso en particular debe pensarse cuál es la variable que es más necesario controlar su influencia sobre los resultados del experimento y buscar aparear a los grupos en esa variable.

El segundo paso consiste en obtener una medición de la variable elegida para emparejar a los grupos. Esta medición puede existir o puede efectuarse antes del experimento. Vamos a suponer que nuestro experimento fuera sobre métodos de enseñanza, el emparejamiento podría hacerse sobre la base de la inteligencia. Si fueran adolescentes, podrían obtenerse —en sus escuelas— registros de inteligencia de ellos o aplicarles una prueba de inteligencia.

El tercer paso consiste en ordenar a los sujetos en la variable sobre la cual se va a efectuar el emparejamiento (de las puntuaciones más altas a las más bajas). Por ejemplo,

supóngase que se tuvieran 16 personas (recuérdese la sugerencia de tener 15 o más en cada grupo, aquí se incluyen 16 únicamente para no hacer muy largo el ejemplo), se ordenarían de la siguiente manera:

SUJETO	COEFICIENTE DE INTELIGENCIA (CI)	SUJETO	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

El cuarto paso es formar parejas de sujetos según la variable de apareamiento (las parejas son sujetos que tienen la misma puntuación en la variable o una puntuación similar) e ir asignando a cada integrante de cada pareja a los grupos del experimento, buscando un balance entre dichos grupos. Supóngase que se tuvieran dos grupos.

Sujeto 01 (CI = 129)

Sujeto 02 (CI = 127)

Grupo 1

Grupo 2

Sujeto 03 (CI = 119)

Sujeto 04 (CI = 119)

Grupo 1

Grupo 2

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

Sujeto 05 (CI = 117)

Sujeto 06 (CI = 116)

Grupo 1

Grupo 2

Sujeto 07 (CI = 114)

Sujeto 08 (CI = 113)

Grupo 1

Grupo 2

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

Sujeto 09 (CI = 110)

Sujeto 10 (CI = 110)

Grupo 1
↓

Grupo 2

Sujeto 11 (CI = 108)

Sujeto 12 (CI = 107)

Grupo 1

Grupo 2

Sujeto 13 (CI = 106)

Sujeto 14 (CI = 105)

↓
Grupo 1

↓
Grupo 2

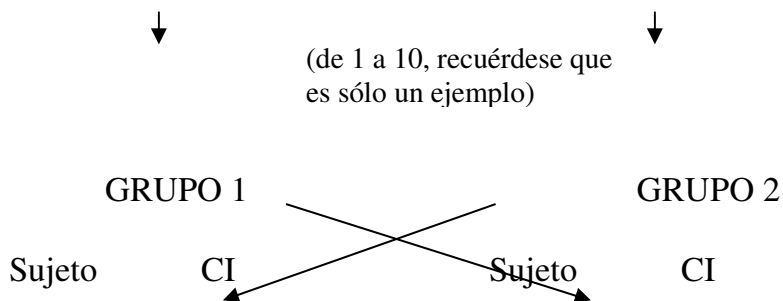
Sujeto 15 (CI = 104)

Sujeto 16 (CI = 102)

↓
Grupo 1

↓
Grupo 2

Los grupos quedarían emparejados en inteligencia.



Cátedra “Metodología para la investigación en Ciencia Política”. Unidad IV
 Hernández Sampieri, R. Collado, L. Lucio, P. Metodología de la investigación (MacGraw Hill, México)

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

Son grupos equivalentes en cuanto a la variable deseada. Este procedimiento puede extenderse a más de dos grupos.

También, podría intentarse emparejar a los grupos en dos variables, pero ambas variables deben estar relacionadas, porque de lo contrario puede resultar muy difícil el emparejamiento. Veámoslo con un ejemplo. Si se deseara aparear a dos grupos por aptitud física y memoria (y digamos que se tuvieran 12 sujetos), podría resultar lo siguiente:

<i>SUJETO</i>	<i>APTITUD FÍSICA</i>	<i>MEMORIA</i>
01	10	1
02	10	6
03	9	9
04	8	2
05	7	4
06	6	10
07	5	2
08	5	9
09	4	3
10	3	1
11	2	10
12	2	6

Y el emparejamiento ya no sería tan exacto. En diversas ocasiones —incluso— no se pueden emparejar los grupos. Asimismo, conforme se tienen más grupos y variables sobre las cuales se pretende aparear a los grupos, el proceso se complica y resulta más difícil poder emparejarlos.

La asignación al azar es la técnica ideal para lograr la equivalencia inicial

Como método para hacer equivalentes a los grupos es superior (más preciso y confiable) la asignación al azar. El emparejamiento no es un sustituto de ésta. El apareamiento puede suprimir o eliminar el posible efecto de la variable apareada, pero no nos garantiza que otras variables (no apareadas) no vayan a afectar los resultados del experimento. En cambio, la aleatorización sí nos garantiza que otras variables (además de la variable independiente o variables independientes de interés para el investigador) no van a afectar a la(s) dependiente(s) y no van a confundir al experimentador. Como comenta Nunnally (1975), la bondad de la asignación al azar de los sujetos a los grupos de un diseño experimental es que el procedimiento garantiza absolutamente que los sujetos no diferirán —en el promedio— con respecto a cualquier característica más de lo que pudiera esperarse por pura casualidad, antes de que participen en los tratamientos experimentales.

6.10. UNA TIPOLOGÍA SOBRE LOS DISEÑOS EXPERIMENTALES GENERALES

En este capítulo se presentarán los *diseños experimentales más comúnmente citados en la literatura sobre experimentación*. Para ello nos basaremos en la tipología de Campbell y Stanley (1966), quienes dividen a los diseños experimentales en tres: a) preexperimentos, b) experimentos “verdaderos” y c) cuasiexperimentos. Asimismo, se utilizará la simbología que normalmente se ha usado en los textos de experimentos.

Simbología de los diseños experimentales

R = *Asignación al azar o aleatorización*. Cuando aparece quiere decir que los sujetos han sido asignados a un grupo de manera aleatoria (proviene del inglés “randomization”).

G = *Grupo de sujetos* (G₁, grupo uno; G₂, grupo dos; etcétera).

X = *Tratamiento, estímulo o condición experimental* (presencia de algún nivel de la variable independiente).

O = *Una medición a los sujetos de un grupo* (una prueba, cuestionario, observación, tarea, etcétera). Si aparece antes del estímulo o tratamiento se trata de una preprueba (previa al tratamiento). Si aparece después del estímulo se trata de una postprueba (posterior al tratamiento).

— *Ausencia de estímulo* (nivel “cero” en la variable independiente). Indica que se trata de un grupo de control.

Asimismo cabe mencionar que la secuencia horizontal indica tiempos distintos y que cuando —en dos grupos— dos símbolos aparecen alineados verticalmente, esto señala que tienen lugar en el mismo momento del experimento. Veamos gráficamente estas dos observaciones.

RG₁

0

X

0

Primero, se asigna a los sujetos al azar, al grupo 1.	Segundo, se aplica una medición previa.	Tercero, se administra el estímulo.	Cuarto, se aplica una medición posterior.
---	---	-------------------------------------	---

RG ₁	X	0	
-----------------	---	---	--

RG ₂	—	0	
-----------------	---	---	--

Ambos símbolos están alineados verticalmente, lo que implica que tienen lugar en el mismo momento.

6.11. 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:

G	X	0
---	---	---

Consiste en administrar un estímulo o tratamiento (una película, un discurso, un método educativo, un comercial televisivo, etcétera) a un grupo, y después aplicar una medición en una o más variables para observar cuál es el nivel del grupo en estas variables.

Este diseño no cumple con los requisitos de un verdadero experimento. 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 — antes del estímulo— el nivel que tenía el grupo en la variable dependiente, ni grupo de comparación. El diseño adolece de los defectos que fueron mencionados al hablar de uno de los requisitos para lograr el control experimental. Ese requisito era tener varios grupos de comparación. No se puede con seguridad establecer causalidad. No controla las fuentes de invalidación interna.

2. Diseño de preprueba-postprueba con un solo grupo

Este segundo diseño se puede diagramar así:

G	0 ₁	X	0 ₂
---	----------------	---	----------------

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

tratamiento.

El diseño ofrece una ventaja sobre el anterior, hay 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 científicos. No hay manipulación ni grupo de comparación y —además— varias fuentes de invalidación interna pueden actuar. Por ejemplo, la historia. Entre O_1 y O_2 pueden ocurrir muchos otros acontecimientos capaces de generar cambios, además del tratamiento experimental (Campbell y Stanley, 1966), y entre mayor sea el lapso entre ambas mediciones mayor será también la posibilidad de que actúe la historia. Asimismo, entre O_1 y O_2 puede presentarse la maduración (fatiga, aburrimiento, etc.).

Por otro lado, se corre el riesgo de elegir a un grupo atípico o que en el momento del experimento no se encuentre en su estado normal. Puede presentarse la regresión estadística y diversas interacciones que se mencionaron (interacción entre selección y maduración, por ejemplo). Asimismo, puede haber un efecto de la preprueba sobre la postprueba. En este segundo diseño tampoco se puede establecer con certeza la causalidad.

Los dos diseños preexperimentales no son adecuados para el establecimiento de relaciones entre la variable independiente y la variable dependiente o dependientes. Son diseños que han recibido bastante crítica en la literatura experimental, y con justa razón porque son débiles en cuanto a la posibilidad de control y validez interna. Los autores de este libro consideramos que su uso debe restringirse a que sirvan como ensayos de otros experimentos con mayor control. Es decir, si alguien piensa efectuar un experimento en forma (utilizando alguno de los diseños que se verán en los apartados “experimentos ‘verdaderos’” y “cuasiexperimentos”), y tiene algunas dudas sobre el estímulo o la manera de administrar las mediciones (por ejemplo, cómo reaccionarán los sujetos al estímulo, cuánto tiempo pueden concentrarse en el experimento o cómo debe ser dada una instrucción); puede —primero— ensayar el experimento mediante un diseño preexperimental (hacer una prueba piloto) y —después— llevar a cabo su experimento utilizando un diseño más confiable.

Asimismo, *en ciertas ocasiones los diseños preexperimentales pueden servir como estudios exploratorios*, pero sus resultados deben observarse con precaución. De ellos no pueden derivarse conclusiones que aseveremos con seguridad. Son útiles como un primer acercamiento con el problema de investigación en la realidad, pero no como el único y definitivo acercamiento con dicho problema. Abren el camino, pero de ellos deben derivarse estudios más profundos.

Desafortunadamente en la investigación comercial, los diseños preexperimentales se utilizan con mayor frecuencia de la que fuera deseable. Algunos investigadores de mercado —por ejemplo— toman a un grupo, lo exponen a un comercial televisivo y miden la aceptación del producto o la predisposición de compra, si ésta es elevada deducen que se debió al comercial. Lo mismo ocurre con programas de desarrollo organizacional,

introducción de innovaciones, métodos de enseñanza, etcétera. Se hacen juicios aventurados y afirmaciones superficiales.

6.12. EXPERIMENTOS “VERDADEROS”²⁴

Los experimentos “verdaderos” son aquellos que reúnen los dos requisitos para lograr el control y la validez interna: 1) grupos de comparación (manipulación de la variable independiente o de varias independientes) y 2) equivalencia de los grupos. Los diseños “auténticamente” experimentales pueden abarcar una o más variables independientes y una o más dependientes. Asimismo, pueden utilizar prepruebas y postpruebas para analizar la evolución de los grupos antes y después del tratamiento experimental. Desde luego, no todos los diseños experimentales utilizan preprueba, pero la postprueba es necesaria para determinar los efectos de las condiciones experimentales (Wiersma, 1986).

1. Diseño con postprueba 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 son asignados a los grupos de manera aleatoria. Después de que concluye el periodo experimental, a ambos grupos se les administra una medición sobre la variable dependiente en estudio. El diseño puede diagramarse de la siguiente manera:

R G ₁	X	O ₁
R G ₂	—	O ₂

En este diseño, la única diferencia entre los grupos debe ser la presencia-ausencia de la variable independiente. Éstos son inicialmente equivalentes y para asegurar que durante el experimento continúen siendo equivalentes —salvo por la presencia o ausencia de dicha manipulación— el experimentador debe observar que no ocurra algo que afecte sólo a un grupo. Debe recordarse que la hora en que se efectúa el experimento debe ser la misma para ambos grupos (o ir mezclando a un sujeto de un grupo con un sujeto del otro grupo —cuando la participación es individual—), lo mismo que las condiciones ambientales y demás factores que fueron ventilados al hablar de equivalencia de grupos.

²⁴ Los autores desean subrayar la importante contribución que para este apartado tuvo la obra de Wierma (1986).

Wiersma (1986), comenta que la postprueba debe ser —preferentemente— administrada inmediatamente después de que concluya el experimento, especialmente si la variable dependiente tiende a cambiar con el paso del tiempo. La postprueba es aplicada simultáneamente a ambos grupos.

La comparación entre las postpruebas de ambos grupos (O_1 y O_2) nos indica si hubo o no efecto de la manipulación. Si ambas difieren significativamente²⁵ ($O_1 \neq O_2$) esto nos indica que el tratamiento experimental tuvo un efecto a considerar. Si no hay diferencias ($O_1 = O_2$), ello nos indica que no hubo un efecto significativo del tratamiento experimental (X). 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 el investigador hipotetiza 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 prejuicio del grupo experimental deberá ser menor que el del grupo de control, $O_1 < O_2$). Pero si O_1 y O_2 son iguales, quiere decir que dicho programa no reduce el prejuicio. Asimismo, puede suceder—incluso— que los resultados vayan en contra de la hipótesis. Por ejemplo, en el caso del prejuicio, si O_2 es menor que O_1 (el nivel del prejuicio, es menor en el grupo que no recibió el tratamiento experimental, el que no vio el programa televisivo).

La prueba estadística que suele utilizarse en este diseño para comparar a los grupos es la *prueba “t”* para grupos correlacionados, al nivel de medición por intervalos.

El diseño con postprueba únicamente y grupo de control puede extenderse para incluir más de dos grupos (tener varios niveles de manipulación de la variable independiente). En este caso, se usan dos o más tratamientos experimentales —además del grupo de control—. Los sujetos son asignados al azar a los distintos grupos, y los efectos de los tratamientos experimentales pueden investigarse comparando las post-pruebas de los grupos. Su formato general sería:

²⁵ Los estudiantes frecuentemente se preguntan: ¿qué es una diferencia significativa? Si el promedio en la postprueba de un grupo en alguna variable es 10 —por ejemplo—, y en el otro es de 12—por ejemplo—; ¿esta diferencia es o no significativa?, ¿puede o no decirse que el tratamiento tuvo un efecto sobre la variable dependiente?

A este respecto, cabe decir que existen pruebas estadísticas que nos dicen si una diferencia entre dos o más cifras (promedios, porcentajes, puntuaciones totales, etcétera) es o no significativa. Estas pruebas toman en cuenta varios aspectos como el tamaño de los grupos cuyos valores se comparan, las diferencias entre quienes integran los grupos (diferencias intragrupo) y otros factores. Cada comparación entre grupos es distinta y esto lo toman en cuenta dichas pruebas, las cuales serán explicadas en el capítulo correspondiente al análisis de los datos. En el presente apartado se mencionarán las pruebas estadísticas más comunes para cada diseño, pero no se explicarán. Los lectores que no estén familiarizados con ellas no tienen —por ahora— que preocuparse de saber qué prueba se utiliza en cada diseño. Para analizar las aplicaciones de las distintas pruebas estadísticas de diferencias de grupos a la experimentación (Cap. 10). No resultaría conveniente explicar las pruebas aquí, porque habría que explicar algunos aspectos estadísticos en los cuales se basan tales pruebas, lo que podría provocar confusión sobre todo entre quienes se inician en el estudio de la investigación del comportamiento humano.

R G ₁	X ₁	O ₁
R G ₂	X ₂	O ₂
R G ₃	X ₃	O ₃
·	·	·
·	·	·

Cátedra “Metodología para la investigación en Ciencia Política”. Unidad IV
Hernández Sampieri, R. Collado, L. Lucio, P. Metodología de la investigación (MacGraw Hill, México)

· R G K	· X _k	· 0 _k
R G K+1	—	0 _{k+1}

Obsérvese que el último grupo no se expone a la variable independiente, es el grupo de control. Si no se tiene grupo de control, el diseño puede llamarse “diseño con grupos aleatorizados y postprueba únicamente” (Wiersma, 1986).

En el diseño con postprueba únicamente y grupo de control, así como en sus posibles variaciones y extensiones, se logra controlar todas las fuentes de invalidación interna. La administración de pruebas no se presenta porque no hay preprueba, al haber una sola medición no puede haber un efecto de la aplicación de una prueba sobre puntuaciones de pruebas subsecuentes (no hay tal). La inestabilidad no afecta porque los componentes del experimento son los mismos para todos los grupos (excepto la manipulación o tratamientos experimentales) ni la instrumentación porque es la misma postprueba para todos, ni la maduración porque la asignación es al azar (si hay —por ejemplo— cinco sujetos en un grupo que se cansan fácilmente, habrá otros tantos en el otro grupo u otros grupos), ni la regresión estadística porque si un grupo está regresando a su estado normal, el otro u otros también. La selección tampoco es problema, porque si hay sujetos atípicos en un grupo, en el otro u otros también habrá sujetos atípicos, y suponiendo que todos en un grupo sean atípicos, en el otro u otros grupos todos serán atípicos y queda compensado (las diferencias se pueden atribuir a la manipulación de la variable independiente y no a que los sujetos sean atípicos, pues la aleatorización hace equivalentes a los grupos en este factor. Si en los dos grupos hubiera solamente personas sumamente inteligentes y la variable independiente es el método de enseñanza, diferencias en aprendizaje pueden atribuirse al método y no a la inteligencia). La mortalidad no afecta, puesto que al ser los grupos equivalentes, el número de personas que abandonen cada grupo tenderá a ser el mismo (salvo que la condición experimental o condiciones tengan algo en especial que haga que los sujetos abandonen el experimento —por ejemplo que las condiciones sean amenazantes para los participantes—. En cuyo caso la situación se detecta y analiza a fondo, de todas maneras el experimentador tiene control sobre la situación debido a que sabe que todo es igual para los grupos con excepción del tratamiento experimental y puede saber por qué se presentó la mortalidad). Otras interacciones tampoco pueden afectar los resultados, pues —por ejemplo— si la selección se controla, sus interacciones operarán de modo similar en todos los grupos. Y la historia se controla si se vigila cuidadosamente que ningún acontecimiento afecte a un solo grupo, si ocurre el acontecimiento en todos los grupos, aunque afecte, lo hará de manera pareja en éstos.

En resumen, lo que influya en un grupo también influirá de manera equivalente en los demás. Este razonamiento se aplica a todos los diseños experimentales “verdaderos”.

Ejemplo del diseño con postprueba únicamente, varios grupos y uno de control

Un investigador lleva a cabo un experimento para analizar cómo influye el tipo de liderazgo que ejerza el supervisor sobre la productividad de los trabajadores.

Pregunta de investigación: *¿Influye el tipo de liderazgo que ejerzan los supervisores de producción —en una maquiladora— la productividad de los trabajadores de línea?*

Hipótesis de investigación: *Distintos tipos de liderazgo que ejerzan los supervisores de producción tendrán diferentes efectos sobre la productividad*”

Hipótesis estadística: $\bar{X}_1 \neq \bar{X}_2 \neq \bar{X}_3 \neq \bar{X}_4$

(\bar{X} = promedios de productividad).

Noventa trabajadores de línea de una planta maquiladora son asignados al azar a tres diferentes condiciones experimentales: 1) 30 trabajadores realizan una tarea bajo el mando de un supervisor con rol autocrático, 2) 30 realizan la tarea bajo el mando de un supervisor con rol democrático, 3) 30 efectúan la tarea bajo el mando de un supervisor con rol “laissez-faire” (que no supervisa directamente, no ejerce presión, es permisivo y desorganizado).²⁶ Finalmente, 30 más son asignados aleatoriamente al grupo de control, donde no hay supervisor. En total, 120 trabajadores.

Se forman grupos de 10 trabajadores para la ejecución de la tarea (armar un sistema de arneses o cables para vehículos automotores). Por lo tanto, habrá 12 grupos de trabajo (repartidos en tres tratamientos experimentales y un grupo de control). La tarea es la misma para todos y los instrumentos de trabajo también, al igual que el ambiente físico de trabajo (iluminación, temperatura, etcétera). Las instrucciones son iguales.

Se ha instruido a tres supervisores (desconocidos para todos los trabajadores participantes) para que puedan ejercer los tres roles (democrático, autocrático y laissez-faire). Los supervisores se distribuyen al azar entre los horarios.

²⁶ Adjetivos utilizados por Sessoms Y Stevenson (1981) para calificar a este tipo de liderazgo.

SUPERVISOR		ROLES	
Supervisor 1	autocrático	democrático	laissez-faire
trabaja a...	10 sujetos	10 sujetos	10 sujetos

Cátedra “Metodología para la investigación en Ciencia Política”. Unidad IV
 Hernández Sampieri, R. Collado, L. Lucio, P. Metodología de la investigación (MacGraw Hill, México)

	(10:00 a 14:00 hrs., el lunes)	(15:00 a 19:00 hrs., el lunes)	(10:00 a 14:00 hrs., el martes)
Supervisor 2 trabaja a...	10 sujetos (15:00 a 19:00 hrs., el lunes)	10 sujetos (10:00 a 14:00 hrs., el martes)	10 sujetos (10:00 a 14:00 hrs., el lunes)
Supervisor 3 trabaja a...	10 sujetos (10:00 a 14:00 hrs., el martes)	10 sujetos (10:00 a 14:00 hrs., el lunes)	10 sujetos (15:00 a 19:00 hrs., el lunes)
Sin supervisor	10 sujetos (10:00 a 14:00 hrs., el lunes)	10 sujetos (15:00 a 19:00 hrs., el lunes)	10 sujetos (10:00 a 14:00 hrs., el martes).

Si se observa, los tres supervisores interactúan en todas las condiciones (ejercen los tres roles), ello con el propósito de evitar que la apariencia física o la personalidad del supervisor influya en los resultados. Es decir, si un supervisor es —por decir— más “carismático” que los demás e influye en la productividad, influirá en los tres grupos.²⁷

La hora está controlada puesto que los tres roles se aplican a todas las horas en que se lleva a cabo el experimento (10:00 a 14:00 hrs y 15:00 a 19:00 hrs el lunes, y 10:00 a 14:00 hrs el martes). Es decir, siempre las tres condiciones están realizándose simultáneamente.

Este ejemplo podría esquematizarse de la siguiente manera:

RG ₁	X ₁	(supervisión con rol autocrático)	0 ₁	
RG ₂	X ₂	(supervisión con rol democrático)	0 ₂	comparaciones en productividad
RG ₃	X ₃	(supervisión con rol laissez-faire)	0 ₃	
RG ₄	—	(sin supervisión)	0 ₄	

Cada tratamiento (X) es aplicado a tres grupos de trabajo y se ha evitado que un solo supervisor intervenga en un único grupo.

²⁷ Alguien podría argumentar que es factible que se presente una interacción entre el supervisor y el rol que adopta, esto es —por ejemplo—, que el supervisor 2 cuando es democrático “se vuelve mas carismático” o que el supervisor 1 cuando es autocrático alcanza más carisma” y ello eleva aún más la productividad. Esto es cierto, pero puede averiguarse mediante métodos estadísticos de análisis y así lograr control, pues éste en última instancia reside en saber qué ocurre con los resultados. Si se presentará este caso, el investigador podría convertir este diseño —a posteriori— en factorial y analizar lo que pasa. Más adelante se discutirán los diseños factoriales. Otra solución sería duplicar los sujetos y que cada supervisor tuviera a dos grupos en cada condición experimental. Ello ayudaría a la interpretación.

2. Diseño con preprueba-postprueba y grupo de control

Este diseño *incorpora la administración de prepruebas a los grupos* que componen el experimento. Los sujetos son asignados al azar a los grupos, después a éstos se les administra simultáneamente la preprueba, un grupo recibe el tratamiento experimental y otro no (es el grupo de control); y finalmente se les administra —también simultáneamente— una postprueba. El diseño puede diagramarse como sigue:

RG ₁	0 ₁	X	0 ₂
RG ₂	0 ₃	—	0 ₄

La adición de la preprueba ofrece dos ventajas: Primera, las puntuaciones de las prepruebas pueden usarse para fines de control en el experimento, al compararse las prepruebas de los grupos se puede evaluar qué tan adecuada fue la aleatorización. Lo cual es conveniente con grupos pequeños. En grupos grandes la aleatorización 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 se puede analizar el puntaje ganancia de cada grupo (la diferencia entre las puntuaciones de la preprueba y la postprueba).

El diseño controla todas las fuentes de invalidación interna por las mismas razones que se argumentaron en el diseño anterior (diseño con postprueba únicamente y grupo de control). Y la administración de la prueba queda controlada, ya que si la preprueba afecta las puntuaciones de la postprueba lo hará similarmente en ambos grupos, 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 de los grupos. En algunos casos, para no repetir exactamente la misma prueba, se pueden desarrollar dos pruebas que no sean las mismas pero que sí sean equivalentes (que produzcan los mismos resultados)²⁸ La historia se contraía observando que ningún acontecimiento solamente afecte al grupo.

Este diseño puede ser extendido para incluir más de dos grupos. Esto podría diagramarse de una manera general del siguiente modo:²⁹

R G ₁	0 ₁		X ₁	0 ₂	
	R G ₂	0 ₃		X ₂	0 ₄
	R G ₃	0 ₅		X ₃	0 ₆

²⁸ Hay procedimientos para obtener pruebas “paralelas” o gemelas”, éstos nos garantizan la equivalencia y se hablará de ellos en el capítulo ‘Recolección de los datos’. 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”.

²⁹ Tomado de Wiersma (1986, p. 111).

Cátedra “Metodología para la investigación en Ciencia Política”. Unidad IV
 Hernández Sampieri, R. Collado, L. Lucio, P. Metodología de la investigación (MacGraw Hill, México)

R G _K	0 _{2k-1}	X _k	0 _{2k}
R G _{K+1}	0 _{2k+1}	—	0 _{2(k+1)}

Se tienen diversos tratamientos experimentales y un grupo de control. Si éste es excluido, el diseño puede llamarse “diseño de preprueba-postprueba con grupos aleatorizados” (Simon, 1986).

Ejemplo del diseño de preprueba—postprueba con grupo de control

Un investigador desea analizar el efecto de utilizar videos didácticos con canciones para enseñar hábitos higiénicos a los niños en edad preescolar.

Pregunta de investigación: *¿Los videos didácticos musicalizados son más efectivos para enseñar hábitos higiénicos a los niños en edad preescolar, en comparación con otros métodos tradicionales de enseñanza?*

Hipótesis de investigación: ‘Los videos didácticos constituyen un *método* más efectivo de enseñanza de hábitos higiénicos a niños en edad preescolar, que la explicación verbal y los folletos instruccionales’.

Hipótesis estadística: $\bar{N}_1 \neq \bar{N}_2 \neq \bar{N}_3 = \bar{N}_4$

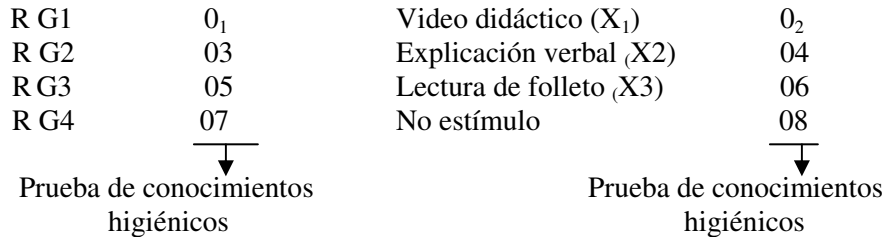
(N = número de hábitos higiénicos aprendidos —en promedio— por cada grupo).

100 niños son asignados al azar a cuatro grupos: 1) un grupo recibirá instrucción sobre hábitos higiénicos por medio de un video con caricaturas y canciones —su duración es de 20 minutos—, 2) este grupo recibirá explicaciones de hábitos higiénicos de una maestra instruida para ello (su explicación durará 20 minutos y no se permiten preguntas. La maestra lleva muy bien preparado lo que debe decir), 3) el tercer grupo leerá un folleto ilustrado con explicaciones sobre hábitos higiénicos (el folleto está diseñado para que un niño — promedio— en edad preescolar, lo lea en 20 minutos), 4) el grupo de control jugará libremente durante 20 minutos. Los grupos permanecerán —simultáneamente— en cuatro salones de clase. Todas las explicaciones contendrán la misma información y las instrucciones son estándares.

Antes del inicio del tratamiento experimental, a todos se les aplicará una prueba sobre conocimiento de hábitos higiénicos —especialmente diseñada para niños—, al igual que inmediatamente después de que hayan recibido la explicación por el medio que les correspondió.

El ejemplo podría esquematizarse de la siguiente forma, como lo muestra la figura 6.3.

FIGURA 6.3
 DIAGRAMA DEL EJEMPLO DE DISEÑO DE PREPRUEBA-POSTPRUEBA,
 TRES GRUPOS EXPERIMENTALES Y UNO DE CONTROL

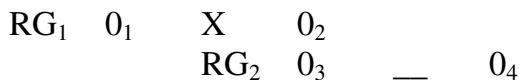


Las posibles comparaciones en este diseño son: a) las prepruebas entre sí (0₁, 0₃, 0₅ y 0₇), b) las postpruebas entre sí para analizar cuál fue el método de enseñanza más efectivo (0₂, 0₄, 0₆ y 0₈) y e) el puntaje-ganancia de cada grupo (0₁ vs. 0₂, 0₃ vs. 0₄, 0₅ vs. 0₆ y 0₇ vs. 0₈), así como los puntajes-ganancia de los grupos entre sí. Y al igual que en todos los diseños experimentales, se puede tener más de una variable dependiente (v.g., terés por los hábitos higiénicos, disfrute del método de enseñanza, etcétera). En este caso, las prepruebas y postpruebas medirían varias dependientes.

Veamos algunos posibles resultados de este ejemplo y sus interpretaciones.

1. Resultado: 0₁ ≠ 0₂, 0₃ ≠ 0₄, 0₅ ≠ 0₆, 0₇ ≠ 0₈; pero 0₂ ≠ 0₄, 0₂ ≠ 0₆, 0₄ ≠ 0₆.
 Interpretación: Hay efectos de todos los tratamientos experimentales, pero son diferentes.
2. Resultado: 0₁ = 0₃ = 0₅ = 0₇ = 0₂ = 0₆ = 0₇ = 0₈, pero 0₃ ≠ 0₄.
 Interpretación: No hay efectos de X₁ y X₃, pero sí hay efectos de X₂.
3. Resultado: 0₁ = 0₃ = 0₅ = 0₇ y 0₂ = 0₄ = 0₆ = 0₈, pero 0₁, 0₃, 0₅ y 0₇ < 0₂, 0₄, y 0₆ y 0₈.
 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).

El análisis estadístico —si se tienen dos grupos— puede ser:



- 1) Para la comparación entre prepruebas se utiliza *la prueba “t”* para grupos correlacionados (nivel de medición por intervalos).
- 2) Lo mismo para la comparación entre las dos postpruebas.
- 3) Igual para analizar —por separado— el puntaje-ganancia de cada grupo (0₁ vs. 0₂ y 0₃ vs. 0₄).
- 4) *Análisis de varianza* (ANOVA) para grupos relacionados si se comparan simul-

táneamente 0_1 , 0_2 , 0_3 y 0_4 y el nivel de medición es por intervalos.

Cuando se tienen más de dos grupos:

1) Para la comparación entre sí de las prepruebas, las postpruebas o todas las mediciones (prepruebas y postpruebas); *el Análisis de Varianza (ANOVA)* para grupos correlacionados, con el nivel de medición por intervalos.

2) Para las mismas comparaciones del punto anterior pero con nivel de medición nominal, *la Ji-cuadrada* para múltiples grupos y *coeficientes para tabulaciones cruzadas*.

3. Diseño de cuatro grupos de Solomon

R. L. Solomon (1949), describió un diseño que era la mezcla de los dos anteriores (diseño con postprueba únicamente y grupo de control más diseño de preprueba-postprueba 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 postprueba. Los sujetos son asignados aleatoriamente. El diseño puede diagramarse así:

R G ₁	0_1	X	0_2
R G ₂	0_3	—	0_4
R G ₃	—	X	0_5
R G ₄	—	—	0_6

El diseño original incluye sólo cuatro grupos y un tratamiento experimental. Los efectos pueden determinarse comparando las cuatro postpruebas. Los grupos 1 y 3 son experimentales, y los grupos 2 y 4 son de control.

La ventaja de este diseño es que el experimentador puede verificar los posibles efectos de la preprueba sobre la postprueba, puesto que a algunos grupos se les administra preprueba y a otros no. Es posible que la preprueba afecte la postprueba o que aquélla interactúe con el tratamiento experimental. Por ejemplo, podría encontrarse lo siguiente — con promedios de una variable determinada— (figura 6.4.):

FIGURA 6.4

EJEMPLO DE EFECTO DE LA PREPRUEBA EN EL DISEÑO DE SOLOMON

R G ₁	$0_1 = 8$	X	$0_2 = 14$
R G ₂	$0_3 = 8.1$	—	$0_4 = 11$
R G ₃	—	X	$0_5 = 11$
R G ₄	—	—	$0_6 = 8$

Teóricamente O_2 debería ser igual a O_5 porque ambos grupos recibieron el mismo tratamiento, igualmente 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 es la preprueba. Las diferencias pueden atribuirse a un efecto de la preprueba (la preprueba afecta —aproximadamente— 3 puntos y el tratamiento experimental también 3 puntos —poco más o menos—). Veámoslo esquemáticamente:

Ganancia con preprueba y tratamiento = 6.

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

Supuestamente —porque la aleatorización hace inicialmente equivalentes a los grupos— la preprueba hubiera sido para todos cerca de ocho, si se hubiera aplicado a todos los grupos. La “supuesta ganancia” (supuesta porque no hubo preprueba) del tercer grupo —con tratamiento y sin preprueba— es de 3. Y la “supuesta ganancia” (supuesta porque tampoco hubo preprueba) del cuarto grupo es nula, 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 8 (calificación que todos —inicialmente— deben tener por efecto de la asignación al azar).

También podría ocurrir un resultado como el de la figura 6.5.

FIGURA 6.5
 EJEMPLO DEL EFECTO DE INTERACCIÓN
 ENTRE LA PREPRUEBA Y EL ESTIMULO EN EL DISEÑO DE SOLOMON

R	G ₁	$O_1 = 7.9$	X	$O_2 = 14$
R	G ₂	$O_3 = 8$	—	$O_4 = 8.1$
R	G ₃	—	X	$O_5 = 11$
R	G ₄	—	—	$O_6 = 7.9$

En el caso de la figura 6.5., la preprueba no afecta (véase la comparación entre O_3 y O_4), 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 postprueba únicamente y grupo de control. La administración de prueba es sometida a análisis minucioso. La historia la controla si se observa que ningún suceso sólo afecte a un grupo.

Las técnicas estadísticas más usuales para comparar las mediciones en este diseño son la

prueba *Ji-cuadrada* para múltiples grupos (nivel de medición nominal), *análisis de varianza* en una sola dirección (Anova Oneway) (si se tiene el nivel de medición por intervalos y se comparan únicamente las postpruebas), y *análisis factorial de varianza* (cuando se tiene un nivel de medición por intervalos y se comparan todas las mediciones —prepruebas y postpruebas—).

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 o estrategias de las psicoterapias. En tales casos, *es conveniente adoptar diseños con varias postpruebas. A estos diseños se les conoce como series cronológicas experimentales.* En realidad el término *serie cronológica*”, *se aplica a cualquier diseño que efectúe — a través del tiempo— varias observaciones o mediciones sobre variable*, sea o no experimental,³⁰ 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 sujetos son asignados al azar a dichos grupos.* Solamente que, debido a que transcurre mucho más tiempo entre el inicio y la terminación del experimento, *el investigador debe tener el suficiente cuidado para 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 al aplicar el estímulo lleva mucho tiempo (por ejemplo, programas motivacionales para trabajadores que pueden durar semanas). Con el paso del tiempo es más difícil mantener la equivalencia inicial de los grupos.

Las series cronológicas experimentales podrían ser diagramadas —por ejemplo— de acuerdo con la figura 6.6.

FIGURA 6.6

EJEMPLOS DE EXPERIMENTOS DE SERIES CRONOLÓGICAS

Serie cronológica sin preprueba, con varias postpruebas y grupo de control

R	G ₁	X ₁	0 ₁	0 ₂	0 ₃
R	G ₂	X ₂	0 ₄	0 ₅	0 ₆
R	G ₃	X ₃	0 ₇	0 ₈	0 ₉
R	G ₄	—	0 ₁₀	0 ₁₁	0 ₁₂

³⁰ En la terminología sobre diseños suelen utilizarse dos términos: “series de tiempo (cronológicas)” y “estudios panel”. Markus (1979) dice que la diferencia principal entre ambos, estriba en que las series de tiempo toman observaciones de un solo ente (individuo, país, empresa, etcétera), 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—.

Cátedra “Metodología para la investigación en Ciencia Política”. Unidad IV
 Hernández Sampieri, R. Collado, L. Lucio, P. Metodología de la investigación (MacGraw Hill, México)

Por su parte, 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, es llamado frecuentemente series múltiples de tiempo o diseño panel de múltiples ondas. En este apartado del libro, los autores han decidido usar el término “series cronológicas múltiples”.

Serie cronológica con preprueba, vanas postpruebas y grupo de control

R	G ₁	0 ₁	X ₁		0 ₂	0 ₃	0 ₄
R	G ₂	0 ₅	X ₂		0 ₆	0 ₇	0 ₈
R	G ₃	0 ₉	—		0 ₁₀	0 ₁₁	0 ₁₂

Serie cronológica basada en el diseño de cuatro grupos de Solomon

R G ₁	0 ₁		X	0 ₂	0 ₃		
R G ₂		0 ₄		—	0 ₅	0 ₆	
R G ₃		—	X	0 ₇	0 ₈		
R G ₄		—	—	0 ₉	0 ₁₀		

Las postpruebas pueden ser tantas como se requieran 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 postpruebas, en cuyo caso se podrían tener 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 está estudiando el efecto que sobre la dedicación en el trabajo pueda tener el hecho de que la directiva de una corporación difunda una serie de valores que aquélla considera que deben ser implantados en ésta.

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”.*

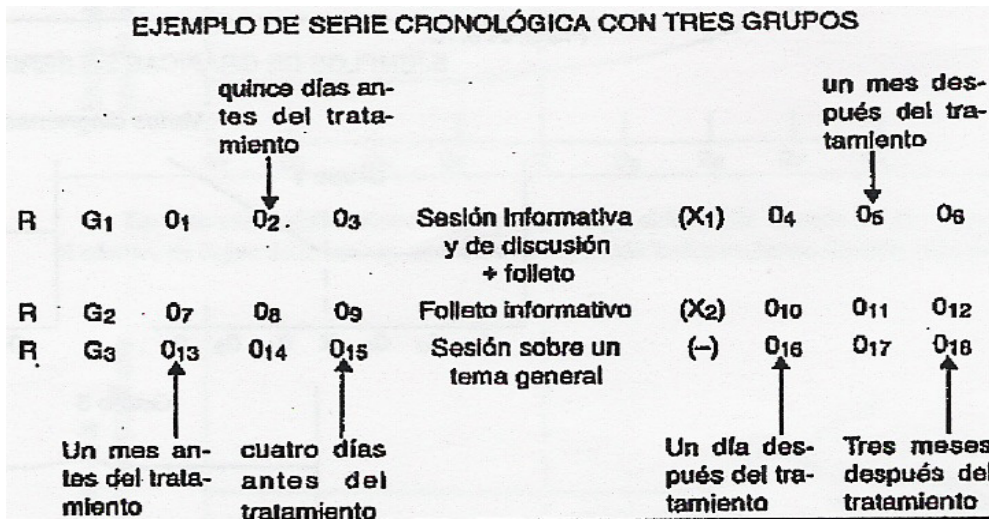
Hipótesis estadística: $r_{xy} > 0$

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 2 horas de duración, en donde se les explican cuidadosamente los valores corporativos —con aplicaciones a situaciones específicas de trabajo—, después se les entrega un folleto con

explicaciones adicionales; 2) otro grupo asiste a una breve sesión donde se proporciona un folleto con explicaciones únicamente sobre los valores corporativos (no hay explicación verbal ni discusión o preguntas y respuestas); 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 generalizado).

Antes de la aplicación de los tratamientos a todos los sujetos se les aplican tres mediciones de la dedicación en el trabajo. Y después de los tratamientos, también se les aplican tres mediciones de la misma variable (al corto, mediano y largo plazo). El diseño podría diagramarse de la siguiente forma (figura 6.7.):

FIGURA 6.7

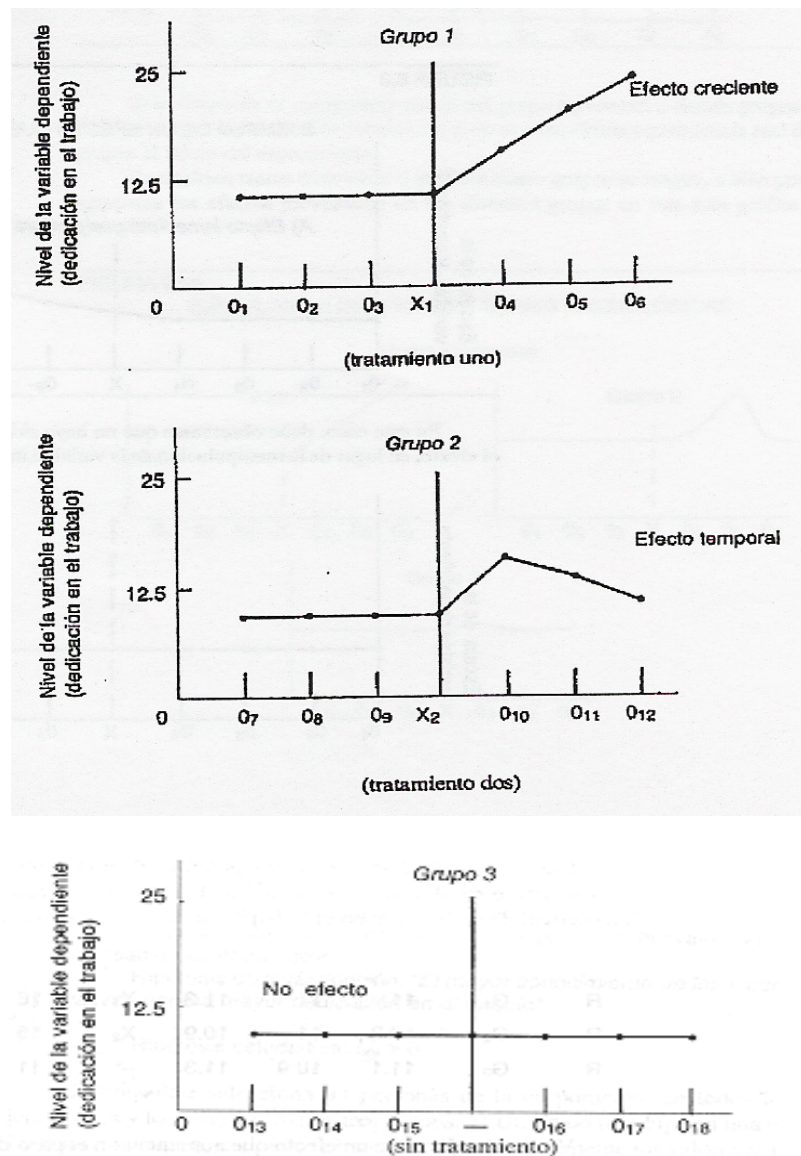


Recuérdese que las mediciones son de la dedicación en el trabajo. Cada grupo estaría integrado por 33 personas.

Una ventaja del diseño es que se puede 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—:

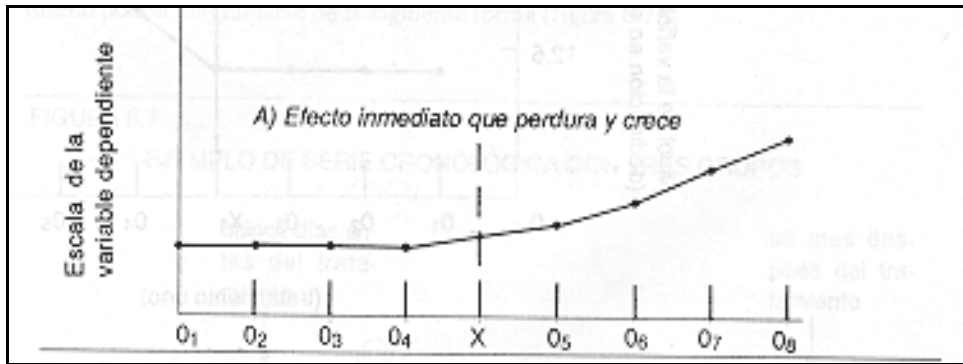
R	G ₁	11	11	11.2	X ₁	16	18	21
R	G ₂	10.8	11	10.9	X ₂	15	14	11.8
R	G ₃	11.1	10.9	11.3	—	11	10.8	11.4

Vemos que X₁ tiene un efecto que aumenta con el paso del tiempo, X₂, produce un efecto en el corto plazo pero éste tiende a desvanecerse con el paso del tiempo. Esto podría graficarse del siguiente modo:



En los diseños experimentales de series cronológicas se puede producir este tipo de gráficas, las cuales enriquecen la interpretación de la evolución de los grupos. Algunas de las diversas configuraciones que pueden presentarse, se muestran en la figura 6.8 y 6.9.

FIGURA 6.8
EJEMPLO DE UN EFECTO QUE PERDURA



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

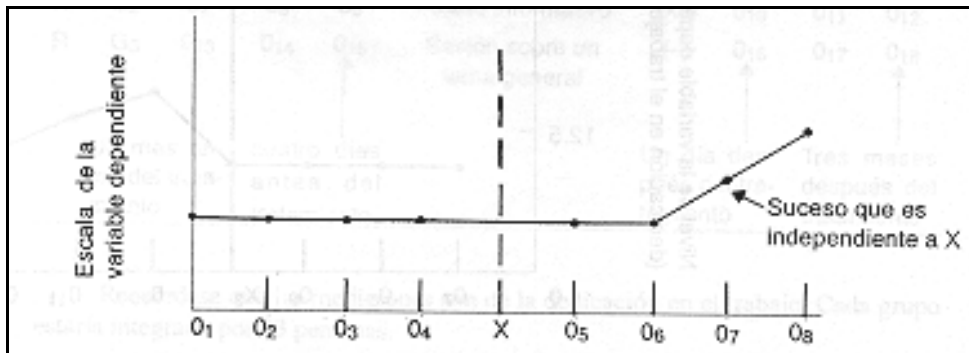
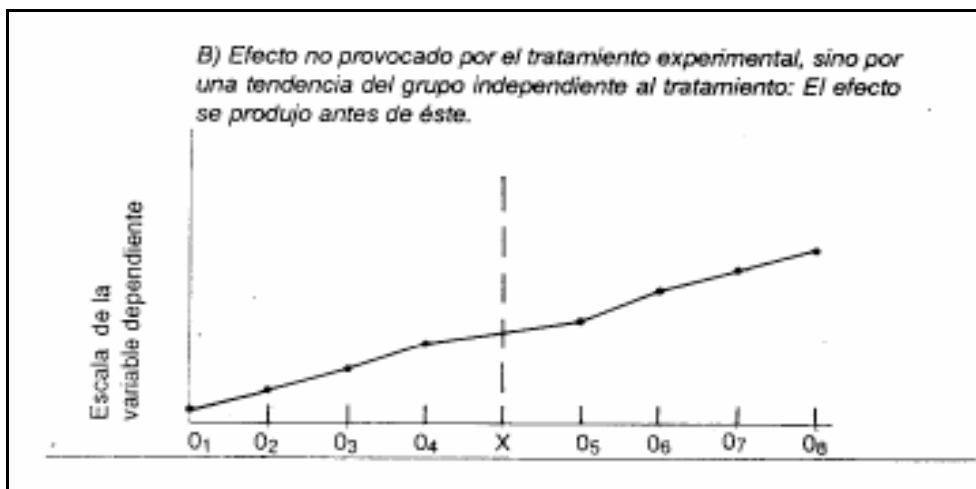


FIGURA 6.9
 EJEMPLO DE AUSENCIA DE EFECTO



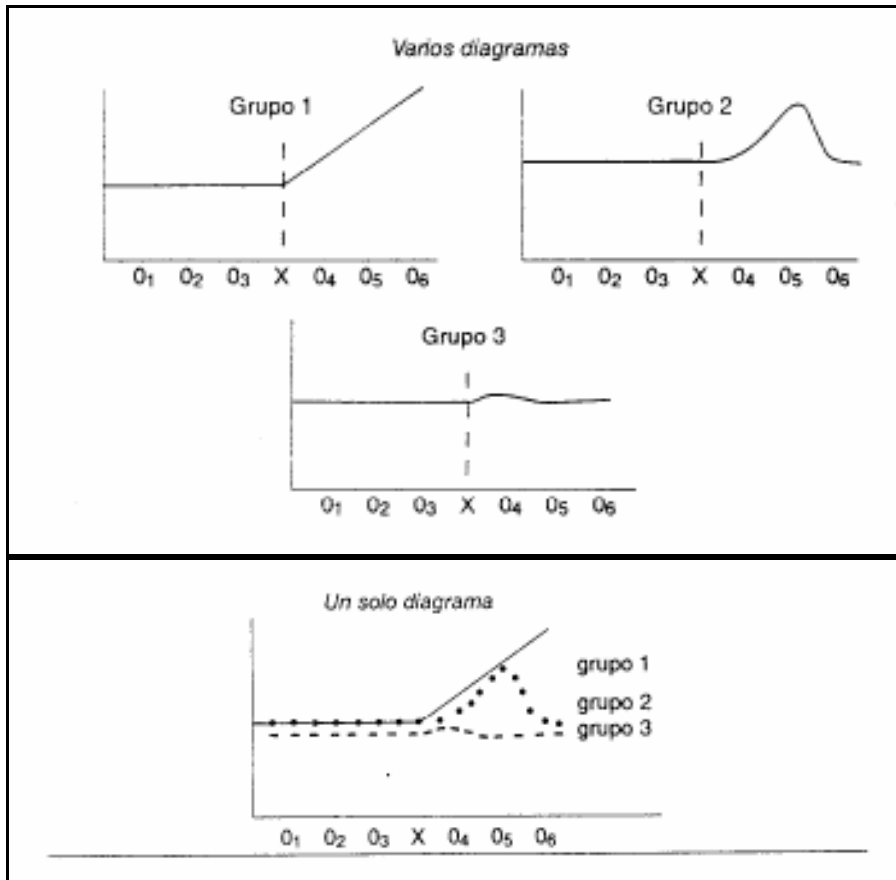
Este diagrama se compararía con el del grupo de control o demás grupos para analizar lo que ocurre. Puede deberse a que no se obtuvo una equivalencia real de los grupos al inicio del

experimento.

Se tendrían tantos diagramas o gráficas como grupos se tengan, o bien podrían agruparse los efectos provocados en los distintos grupos en una sola gráfica. Por ejemplo:

FIGURA 6.10

EJEMPLOS DE GRÁFICAS EN SERIES CRONOLÓGICAS



Desde luego, si se esta midiendo mas de una variable dependiente, Se tendrían en el primer caso un diagrama para cada grupo por cada variable dependiente, y en el segundo caso un diagrama por cada variable dependiente.

EJEMPLO

Si el tratamiento es un programa motivacional y las variables dependientes son la motivación intrínseca, la productividad y la calidad en la producción (teniendo dos grupos experimentales y uno de control), se tendrían las opciones que se muestran en las figuras 6.11. y 6.12.

FIGURA 6.11

EJEMPLO DE VARIOS DIAGRAMAS PARA CADA VARIABLE DEPENDIENTE

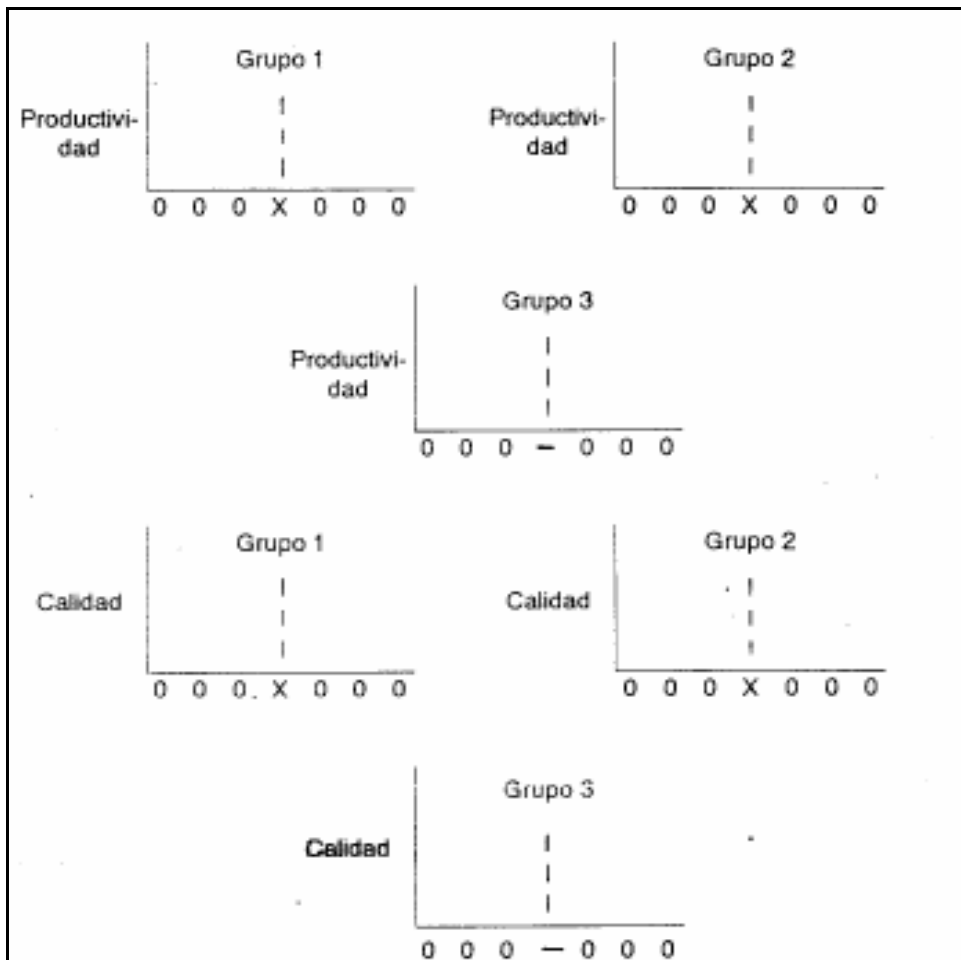
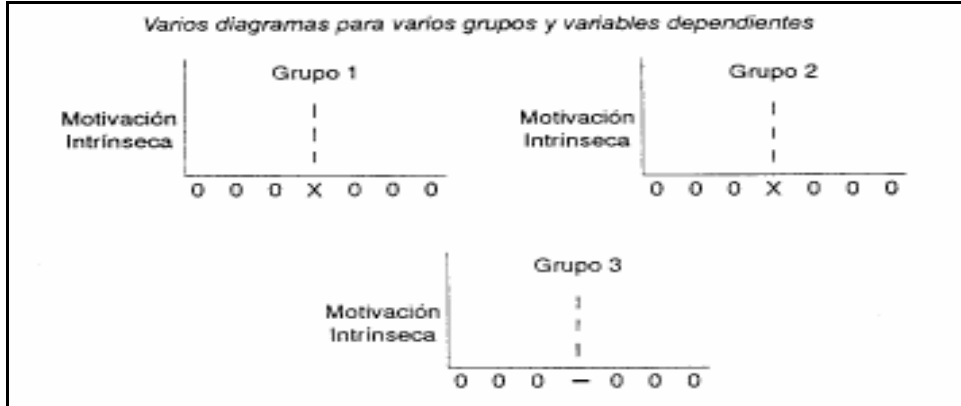
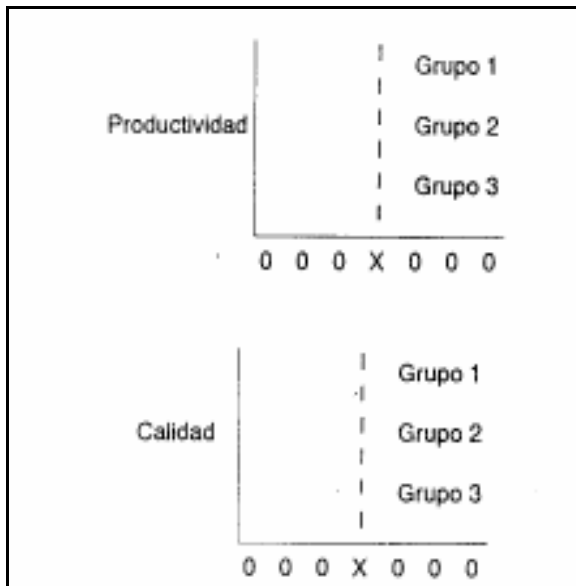
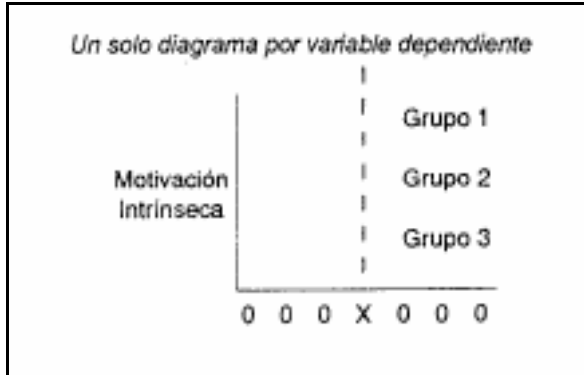


FIGURA 6.12
EJEMPLO DE UN SOLO DIAGRAMA POR VARIABLE DEPENDIENTE



En estos diseños de series cronológicas, se controlan todas las fuentes de invalidación interna, siempre y cuando se lleve a cabo un seguimiento minucioso de los grupos, para asegurarse que la única diferencia entre ellos es la manipulación de la variable independiente. En algunos casos, puede haber una influencia de la repetición de las mediciones sobre la variable dependiente (administración de pruebas múltiples), sobre todo en las pruebas donde el sujeto participa activamente y de manera consciente de que está respondiendo a una prueba (cuestionarios, entrevistas, pruebas estandarizadas), no tanto así en las pruebas en que el sujeto es más pasivo y no se encuentra consciente de que se le mide (v.g., la observación). De cualquier manera, en caso de que se presente dicha influencia, se presentará de manera similar en todos los grupos (porque son equivalentes y el número de pruebas aplicadas es el mismo).

Para estos diseños se suelen utilizar diversas técnicas estadísticas complejas,
- 52 -

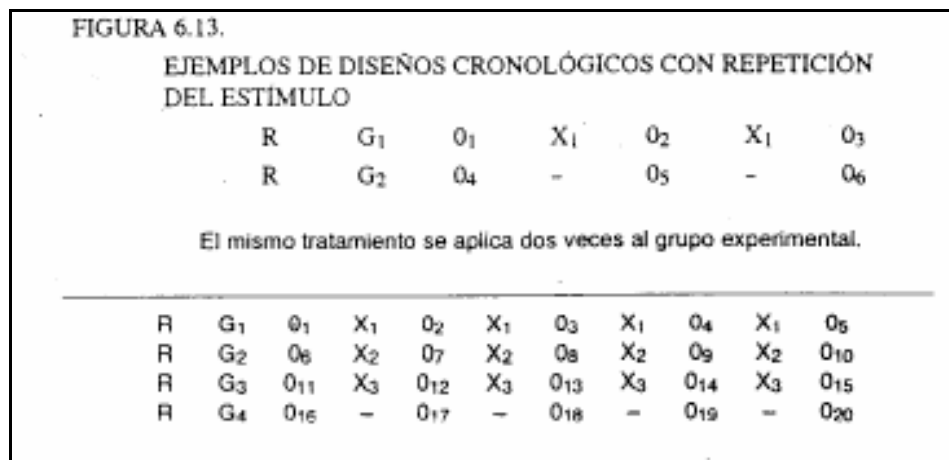
dependiendo del nivel de medición de las variables y el tipo de análisis e interpretación que se desee hacer; tales como análisis de regresión múltiple, análisis de cambio. Se recomiendan las siguientes fuentes para conocer estos análisis: Markus (1979), Ostrom (1978), Kessler y Greenberg (1981), Henkel (1976), Siegel (1980), Mange y Cappella (1980), y Kerlinger y Pedhazur (1973).

5. Diseños de series cronológicas con repetición del estímulo

En ocasiones, el investigador anticipa que el tratamiento a estímulo experimental no tiene efecto o este es mínima si se aplica una sola vez, tal como sería hacer ejercicio físico un solo día, no se puede esperar un cambio en la musculatura; o como sería consumir vitaminas por una única vez. También a veces el investigador quiere conocer el efecto sobre las variables dependientes, cada vez 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 se puede repetir el tratamiento experimental y administrar una posprueba después de cada aplicación de este, para evaluar el efecto de cada aplicación.*

Los sujetos son asignados al azar a los distintos grupos y a cada grupo se le administra varias veces el tratamiento experimental que le corresponde. Algunos de estos diseños diagramados, se muestran en la figura 6.13.



Cada tratamiento se aplica cuatro veces al grupo respectivo.

En algunos casos se podría prescindir de las prepruebas, y el experimentador pudiera –por

alguna justificación teórica o empírica- *aplicar postpruebas a intervalos sistemáticos diferentes*. Por ejemplo:

R	G ₁	X ₁	O ₁	X ₁	X ₁	X ₁	O ₂	X ₁	X ₁	X ₁	O ₃
R	G ₂	X ₂	O ₄	X ₂	X ₂	X ₂	O ₅	X ₂	X ₂	X ₂	O ₆
R	G ₃	-	O ₇	-	-	-	O ₈	-	-	-	O ₉

O bien, aplicar las postpruebas a intervalos irregulares (par determinada razón):

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

Un ejemplo de estos diseños, sena ci caso de un publicista que pretende analizar los efectos de un comercial televisivo sobre —digamos— la preferencia del producto anunciado en relación con otras marcas, y que hipotetiza que una sola exposición al comercial no tendrá efecto.

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

6. Diseños con tratamientos múltiples

A veces, el investigador desca analizar *el efecto de aplicar los diversos tratamientos experimentales a todos los sujetos*. En estos casos se pueden utilizar los diseños con tratamientos múltiples. La aplicación de tratamientos puede ser individual a en un grupo y hay distintas variaciones:

A) *Varios grupos*

En este caso, se tienen varios grupos y los sujetos son asignados al azar a los diferentes grupas, a los cuales se les aplican todos los tratamientos. *La secuencia* de la aplicación de tratamientos *puede ser la misma* para todos los grupos a diferente y se puede administrar una a más postpruebas a los grupos (posteriores a cada tratamiento experimental). Dos diagramas que ejemplifican a estos diseños son los siguientes (figura 6.14.).

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

Con *secuencia diferente*, el experimentador debe tener cuidado al interpretar las segundas postpruebas y mediciones subsecuentes, ya que puede haber 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, y 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 las grupos, así como los dos casos que vamos a ver a continuación (B y C) pueden tener distintos efectos que tienen que analizarse minuciosamente. Algunos tratamientos tienen efectos reversibles, en esta situación no hay interferencia entre tratamientos, y las postpruebas se ven influidas únicamente por el tratamiento inmediato anterior (por ejemplo O₃ —del diseño con secuencia diferente— se vería afectada por X₃, pero no por X₂ a X₁), y ella facilita la interpretación. Pero frecuentemente los efectos no son reversibles, sino aditivas e interactivos; esto es, los resultados de una postprueba se pueden ver influidos no solamente por el tratamiento inmediatamente anterior sino por los que le antecedieron a éste, y no es fácil saber —por ejemplo— cuánto se debió a X₁, cuánto a X₂, a X_k. Para ello, debe incluirse en el análisis el factor secuencia.

B) Un solo grupo

En situaciones donde por algún motivo se cuenta con un número reducido de sujetos para el experimento, se puede llevar a cabo un *diseño con tratamientos múltiples y un solo grupo*. No hay asignación al azar puesta 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 podría diagramarse así:

$$G \text{ único } X_1 \ O_1 \ X_2 \ O_2 \text{ — } O_3 \ X_3 \ O_4 \text{ — } O_5 \ X_k \ O_{k+1} \dots$$

Cuando se considere conveniente, se utiliza como grupo de control, por ejemplo, antes de O₃ y O₅. Sin embargo este diseño, está limitado a que los efectos de los tratamientos múltiples sean reversibles; de lo contrario no es un diseño experimental sino cuasiexperimental. Y 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 este punto se han revisado diseños experimentales que manipulan una sola variable independiente (los tratamientos representan niveles de presencia a manipulación de esta; X_1, X_2, X_3, X_k son variaciones de la misma variable independiente —al menos como han sido concebidas en este libro³¹— y miden una a más variables dependientes (v.g., en un experimento con métodos educativos, en lugar de medir nada más el aprendizaje, se puede medir —además— la motivación del alumno, su integración al grupo, etcétera; y esta en la preprueba y la postprueba). Normalmente, se miden varias variables dependientes, porque si se va a realizar un experimento, con el costo que este implica, bien vale la pena sacarle de una vez el máximo provecho, analizando los efectos de la variable independiente sobre varias dependientes. Además—en ocasiones— como parte de la(s) postprueba(s) se incluyen mediciones para verificar que tan bien funcionó la manipulación (verificaciones de la manipulación, ver página 117), o a veces estas verificaciones son independientes de la(s) postprueba(s). Ahora hablaremos de los *experimentos que incorporan dos a más variables independientes* en el diseño: Los factoriales.

7. 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. Han sido sumamente utilizados en la investigación del comportamiento. La construcción básica de un diseño factorial consiste en que todos los niveles de cada variable independiente son tomados en combinación con todos los niveles de las otras variables independientes (Wiersma, 1986, p. 115).

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:

2	X	2
Un dígito (primera variable independiente)		Otro dígito (segunda variable independiente)

³¹ Algunos autores consideran que cuando se introducen sistemáticamente otros elementos a los diseños tales como presencia-ausencia de la preprueba (v.g., el diseño de cuatro grupos de Solomon), secuencias diferentes de tratamientos en varios grupos para ver si las secuencias se relacionan con características de los grupos (v.g., 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

Cátedra “Metodología para la investigación en Ciencia Política”. Unidad IV
Hernández Sampieri, R. Collado, L. Lucio, P. Metodología de la investigación (MacGraw Hill, México)

G4 X2 07 X1 08
Niños de 4o. año, G5 X1 09 X2 010
G6 X2 011 X1 012),

o el numero de postpruebas diferentes en los grupos (por ejemplo con un diseño como el siguiente:

R	G1	X1	01	02
R	G2	X1	01	
R	G3	X2	01	02
R	G4	X2	01	
R	G5	—	01	02),

se tienen diseños factoriales. Y de hecho, tienen razón porque se están manipulando 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 estos se vean primero.

Y el valor numérico de cada dígito indica el número de niveles de la variable independiente en cuestión. En este caso es “2”, esto quiere decir que cada una de las variables tiene dos niveles. Como menciona Wiersma (1986), 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 $2 \times 2 \times 3$ indica que hay tres variables independientes, la primera y la segunda con dos niveles, mientras que la tercera con tres niveles. El diseño factorial $4 \times 5 \times 2 \times 3$, 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 2×2 sería tener como variables independientes “método de enseñanza y sexo”. 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. Así, en un diseño 2×2 tendremos cuatro grupos ($2 \times 2 = 4$); en un diseño 3×2 tendremos seis grupos; y en un diseño $3 \times 3 \times 3$ tendremos veintisiete grupos. Debe observarse que el resultado de la multiplicación es el número de grupos resultante. En estos diseños, el número de grupos aumenta rápidamente con el incremento del número de variables independientes y/o niveles (exponencialmente).

Veámoslo:

$$2 \times 2 = 4$$

$$2 \times 3 = 6$$

$$3 \times 3 = 9$$

$$3 \times 4 = 12$$

$$3 \times 2 \times 2 = 12$$

$$3 \times 3 \times 3 = 27$$

Ello se debe a que los niveles deben tomarse en todas sus posibles combinaciones

entre sí.

Wiersma (1986) comenta 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 (v.g., sexo, edad, año, escolaridad, inteligencia, etcétera).

Para simplificar la forma en que se diagraman los diseños factoriales, acudiremos a la simbología que comúnmente se utiliza.³² Para designar a las variables independientes se

³² V.g., Matheson, Bruce y Beauchamp (1980); Christensen (1980); Amau-Grass (1981); y Wiersma (1986)

usan letras, (A, B, C, ...K) y para los niveles números (1, 2, 3, ...K), las combinaciones de letras y números que aparecen en las casillas (o celdas) represen tan las mezclas de niveles de las variables independientes. Cada celda es un grupo. En La figura 6.15., se diagrama un diseño factorial 2 x 2.

FIGURA 6.15

DISEÑO 2 X 2

		Variable independiente A	
		A ₁	A ₂
Variable independiente B	B ₁	A ₁ B ₁	A ₂ B ₁
	B ₂	A ₁ B ₂	A ₂ B ₂

Otro ejemplo sería un diseño factorial 2 x 4 x 3 (figura 6.16)

FIGURA 6.16

DISEÑO 2 x 4 x 3

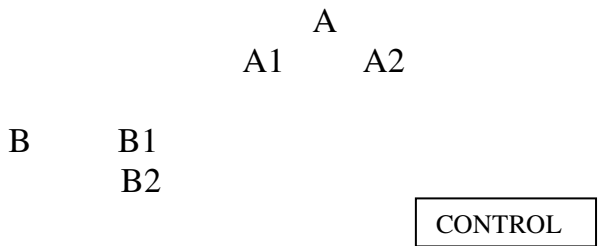
		A					
		A ₁			A ₂		
		C ₁	C ₂	C ₃	C ₁	C ₂	C ₃
B	B ₁	A ₁ B ₁ C ₁	A ₁ B ₁ C ₂	A ₁ B ₁ C ₃	A ₂ B ₁ C ₁	A ₂ B ₁ C ₂	A ₂ B ₁ C ₃
	B ₂	A ₁ B ₂ C ₁	A ₁ B ₂ C ₂	A ₁ B ₂ C ₃	A ₂ B ₂ C ₁	A ₂ B ₂ C ₂	A ₂ B ₂ C ₃
	B ₃	A ₁ B ₃ C ₁	A ₁ B ₃ C ₂	A ₁ B ₃ C ₃	A ₂ B ₃ C ₁	A ₂ B ₃ C ₂	A ₂ B ₃ C ₃
	B ₄	A ₁ B ₄ C ₁	A ₁ B 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

Obsérvese en la figura 6.16., que todas las posibles combinaciones de niveles entre A, B y C están presentes, además ninguna combinación es exactamente igual a la otra. Cada combinación representa una celda o grupo. Si las tres variables habrán de ser manipuladas deliberadamente (al menos una debe de ser para que hablemos de experimento), los sujetos deben ser asignados al azar a todas las celdas o grupos. Si dos variables habrán de ser manipuladas intencionalmente (v.g., B y C), los sujetos de cada nivel de La variable restante

serán asignados al azar a las casillas que les corresponde. Veámoslo con un ejemplo. Si A = Sexo (A1, masculino; A2, femenino), B = Violencia televisada (B4, elevada; B3, mediana; B2, baja y B1, nula) y C = Orientación sobre el programa visto (C1, de los dos padres; C2, del padre y C3, de La madre). Teniéndose 120 niños y 120 niñas, los niños (A1) se asignarían al azar a las celdas en donde A1 esta presente (10 niños en cada celda), y las niñas (A2) a las doce casillas restantes (donde A2 está presente). Si una sola variable es la que se manipula deliberadamente (C por ejemplo), los sujetos de los niveles combinados de las otras dos variables se asignan al azar a los niveles de aquélla (C1, C2 y C3 —en el ejemplo—). Los sujetos A1 B1 serían asignados aleatoriamente a C1, C2 y C3, igual los sujetos A1 B2, A1 B3, etcétera.

A1B1	A2B1
A1B2	A2B2

En los diseños factoriales se puede agregar un grupo de control o varios (que no se expongan a la variable o variables manipuladas deliberadamente. Por ejemplo:



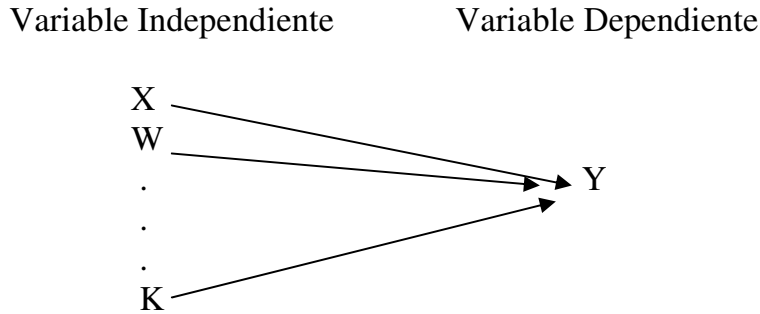
Utilidad de los diseños factoriales

Los diseños factoriales son sumamente útiles porque permiten al investigador evaluar los efectos de cada variable independiente sobre La dependiente por separado y los efectos de las variables independientes conjuntamente. A través de estos diseños se pueden observar los efectos de interacción entre las variables independientes.

En términos de Wiersma (1986, p. 116), La interacción es un efecto producido sobre La variable dependiente, de tal manera que el efecto de una variable independiente deja de permanecer constante a través de los niveles de La otra. 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 si 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 (o viceversa, es sólo un ejemplo ficticio). O el caso de que los métodos de enseñanza aumentan el aprendizaje de las niñas pero no de los niños.

Así, hay dos tipos de efectos que se pueden 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 si y otras dos entre si o pueden interactuar tres o las cuatro variables independientes). Los diseños factoriales responden a estructuras entre variables que podrían esquematizarse en la siguiente forma:



Y analizan el efecto de X sobre Y ($X \rightarrow Y$), el efecto 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 $\left[\begin{matrix} X \\ : \\ W \end{matrix} \right] \rightarrow Y$, el efecto conjunto de X y...K sobre Y $\left[\begin{matrix} X \\ : \\ W \end{matrix} \right] \rightarrow Y$, el efecto conjunto de W y...K sobre Y $\left[\begin{matrix} W \\ : \\ K \end{matrix} \right] \rightarrow Y$, y el efecto conjunto de todas las variables independientes sobre Y $\left[\begin{matrix} X \\ W \\ : \\ K \end{matrix} \right] \rightarrow Y$

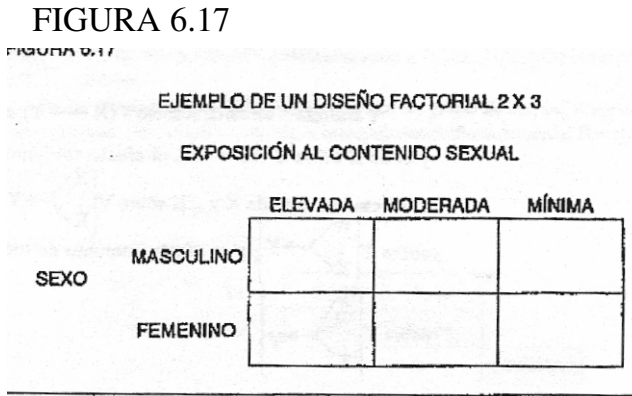
Christensen (1980, pp. 182 y 183), desarrolla una excelente explicación de los efectos principales e interactivos en los diseños factoriales, La cual sirve de base para la exposición de dichos efectos que a continuación se incluye.

UN EJEMPLO DE DISEÑO FACTORIAL

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”. Entonces se diseñará 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: sexo. Entonces se tiene un diseño factorial con dos variables independientes. La exposición tendría tres niveles: a) contenido sexual elevado y manifiesto (por ejemplo, videos donde las modelos femeninas aparecen con ropa muy ligera, hay presencia de roce entre mujeres y hombres y flirteo, se insinúa futuro contacto sexual, etcétera), b) contenido sexual moderado (videos similares a los anteriores, pero donde no se insinúa futuro contacto sexual y las modelos femeninas aparecen con ropa menos ligera) y c) contenido “romántico” (en los videos aparecen las modelos con ropa discreta y las situaciones no sugieren futuro contacto sexual, por el contrario, manifiestan relaciones en el presente sin posibilidad de futuro contacto sexual).

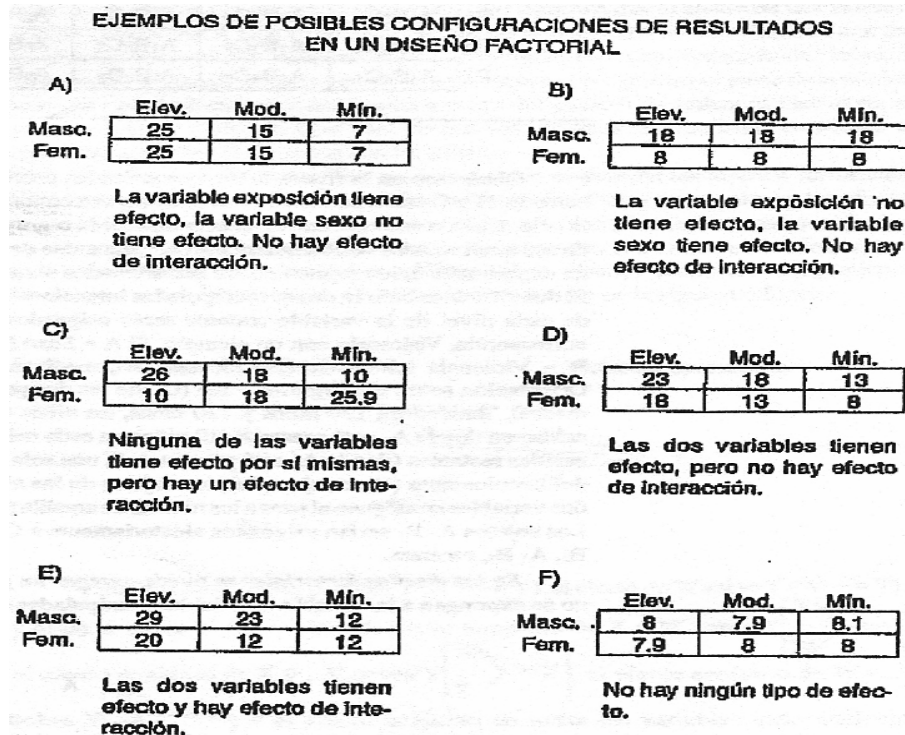
Cada uno de los videos que se elaboren tendrá —entonces— tres versiones. La música es la

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 sexo tendría sus dos niveles: masculino y femenino. El diagrama del diseño se indica en la figura 6.17.

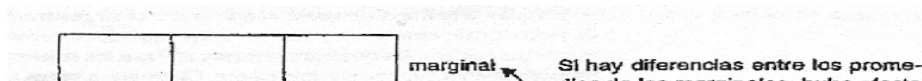


En caso de haber sido expuestos a los videos, se les aplica una postprueba que mide su predisposición para establecer contacto sexual y se obtiene el promedio de cada grupo (asimismo, pensemos que los resultados que pueden obtenerse en esta prueba oscilan entre 0 y 30, donde un valor mayor indica una mas alta predisposición). Analizamos varias configuraciones posibles de resultados:

FIGURA 6.18



En La configuración A, no hay diferencias por sexo (cada casilla en cada nivel de sexo es igual a la del otro nivel). En cambio hay diferencias entre los niveles de exposición



en ambos sexos. Los efectos principales (es decir, los de cada variable por separado) se pueden ver comparando los promedios de los marginales de los niveles de cada variable. Por ejemplo:

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

En la configuración C, las diferencias entre las celdas se explican porque las dos variables interactúan (cuando el sexo es masculino y La exposición elevada, y el sexo es femenino y La exposición mínima, se obtiene un valor; cuando el sexo es masculino y La exposición mínima, y el sexo es femenino y La exposición elevada, se obtiene otro valor; y —finalmente— cuando ambos sexos se exponen moderadamente se obtiene un valor distinto a las demás celdas). 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 solas, 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 en 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 Ji-cuadrada para múltiples grupos —con dicha variable medida

nominalmente—.

Finalmente, a estos diseños se les pueden agregar más variables dependientes (tener dos o mis) y se convierten en diseños multivariados experimentales que utilizan como método estadístico de análisis el Análisis Multivariado de Varianza (MANOVA).

6.13. ¿QUE ES LA VALIDEZ EXTERNA?

Un experimento debe buscar ante todo validez interna, es decir, confianza en los resultados. Si no se logra, no hay experimento verdadero. La primero es eliminar las fuentes que atentan contra dicha validez. Pero la validez interna es sólo una parte de La validez de un experimento, en adición a ella es muy deseable que el experimento tenga validez externa. La validez externa tiene que ver con que’ tan generalizables son los resultados de un experimento a situaciones no experimentales y a otros sujetos o poblaciones. Responde a la pregunta: ¿Lo que encontró en el experimento a que sujetos, poblaciones, contextos, variables y situaciones puede aplicarse?

Por ejemplo, si hacemos un experimento con métodos de aprendizaje y los resultados pueden generalizarse a la enseñanza cotidiana en las escuelas de educación básica, el experimento tendrá validez externa. Si se pueden generalizar a la enseñanza cotidiana en un distrito escolar hay cierta validez externa, si se puede generalizar a la enseñanza cotidiana en un estado, provincia a departamento hay mayor validez externa; y si se pueden generalizar a La enseñanza cotidiana del país y la enseñanza de nivel media, aún mayor validez externa.

Así, resultados de experimentos sobre obediencia a la autoridad que puedan generalizarse a situaciones diarias de trabajo, situaciones familiares, de guerra, etcétera; son experimentos con validez externa.

Fuentes de invalidación externa

Existen diversos factores que pueden amenazar la validez externa, los mas comunes son los siguientes:

1. *Efecto reactivo o de interacción de las pruebas*

Se presenta cuando La preprueba aumenta o disminuye La sensibilidad La calidad de la reacción de los sujetos a la variable experimental, hacienda que los resultados obtenidos para una población con preprueba no puedan generalizarse a quienes forman parte de esa población pero sin preprueba (Campbell y Stanley, 1966). Babbie (1979) utiliza un excelente ejemplo de esta influencia: En un experimenta diseñado para analizar si una película disminuye el prejuicio racial, la preprueba podría sensibilizar al grupo experimental y la película tener un efecto mayor del que tendría si no se aplicara la preprueba (por ejemplo, Si se pasara la película en un cine a en La televisión). Esto es, puede suceder que La película

sólo tenga efecto cuando se administra la preprueba.

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

Este factor se refiere a que se eligen personas con una a varias características que hagan que el tratamiento experimental tenga efecto, el cual no se dan si las personas no tuvieran esas características. Par ejemplo, si seleccionamos trabajadores sumamente motivados para un experimento sobre productividad, podría ocurrir que el tratamiento sólo tuviera efecto en este tipo de trabajadores y no en otros. Ella podría resolverse con una muestra representativa de todos los trabajadores a introduciendo un diseño factorial y una de las variables fuera el grado de motivación.

Este factor puede presentarse en algunos experimentos donde se recluten voluntarios.

3. *Efectos reactivos de los tratamientos experimentales*

La “artificialidad” de las condiciones puede hacer que el contexto experimental resulte atípico respecta a La manera coma se aplica regularmente el tratamiento (Campbell, 1975). Par ejemplo, a causa de La presencia de observadores y equipo los sujetas pueden alterar su conducta normal en La variable dependiente medida, la cual no la alterarían en una situación común donde se aplicara el tratamiento. Par ella, el experimentador debe ingeniárselas para hacer que los sujetos se olviden que están en un experimento y no deben saberse observados.

4. *Interferencia de tratamientos múltiples*

Si los tratamientos no son de efecto reversible, es decir, si no se pueden borrar sus efectos, las conclusiones solamente podrán hacerse extensivas a las personas que experimentaron la misma secuencia de tratamientos, sean múltiples a repetición del misma.

5. *Imposibilidad de replicar los tratamientos*

Cuando los tratamientos son tan complejos que no pueden replicarse en situaciones no experimentales, es difícil generalizar a éstas.

Para lograr una mayor validez externa, es conveniente tener grupos lo más parecidos posible a la mayoría de las personas a quienes se desea generalizar y repetir el experimento varias veces con diferentes grupos (hasta donde el presupuesto y los costos de tiempo lo permitan). También, tratar de que el contexto experimental sea lo más similar posible al

contexto al que se pretende generalizar. Por ejemplo, si SC trata de métodos de enseñanza podría resultar muy conveniente que se usen aulas sumamente parecidas a las que normalmente utilizan los sujetos y que las instrucciones las proporcionen las maestras de siempre. Clara que a veces no es posible. Sin embargo, el experimentador debe esforzarse para que no sientan los participantes —o La sientan La menos que sea factible— que se está experimentando con ellos.

6.14. ¿CUALES PUEDEN SER LOS CONTEXTOS DE EXPERIMENTOS?

En la literatura sobre la investigación del comportamiento se ha distinguido entre dos contextos en donde puede tomar lugar un diseño experimental: Laboratorio y campo. Así, se habla de experimentos de laboratorio y experimentos de campo. Kerlinger (1975, p. 146), define al experimento de laboratorio como: “un estudio de investigación en el que la variancia” (efecto) “de todas a casi todas las variables independientes influyentes posibles no pertinentes al problema inmediato de La investigación se mantiene reducida” (reducida —el efecto—) “en un mínimo”. El mismo autor define el experimento de campo como: ‘un estudio de investigación en una situación realista en La que una a más variables independientes son manipuladas por el experimentador en condiciones tan cuidadosamente controladas como La permite La situación’ (p. 419). La diferencia esencial entre ambos contextos es La “realidad” con que los experimentos se llevan a cabo, el grado en que el ambiente es natural para los sujetos.

Por ejemplo, si creamos salas para ver televisión y las acondicionamos de tal modo que se controle el ruido exterior, la temperatura y otros distractores. Además, incluimos equipo de filmación oculto. Y llevamos —por ejemplo— a los niños para que vean programas de televisión previamente grabados, estamos realizando un experimento de laboratorio (situación creada “artificialmente”). En cambio, si el experimento se lleva a cabo en el ambiente natural de los sujetos (su escuela, fábrica donde trabajan, hogar, etcétera), se trata de un experimento de campo. Pero en ambos casos se lleva a cabo un experimento, siempre y cuando se manipule intencionalmente una variable independiente.

Los experimentos de laboratorio generalmente logran un control más riguroso que los experimentos de campo (Festinger, 1975; Kerlinger 1975), pero estos últimos suelen tener mayor validez externa. Ambos tipos de experimento son deseables.

Algunos han acusado a los experimentos de laboratorio de “artificialidad”, de tener poca validez externa, pero como argumenta Kerlinger (1975): los objetivos primarios de un experimento verdadero son descubrir relaciones (efectos) en condiciones “puras” y no contaminadas, probar predicciones de teorías y refinar teorías e hipótesis. Y comenta: “Realmente, es difícil saber si La artificialidad es una debilidad o simplemente una característica neutral de las situaciones experimentales de laboratorio. Cuando se prepara

deliberadamente una situación de investigación para excluir las muchas distracciones del medio, es quizá ilógico designar a La situación con un término que exprese en parte el resultado que se busca. La crítica de La artificialidad no proviene de las experimentadores, quienes saben que las situaciones experimentales son artificiales; proviene de individuos que carecen de una comprensión de las metas de los experimentas de laboratorio” (p. 417).

Festinger (1975, p. 139) señala (al responder a La crítica de “artificialidad”): “Esta crítica requiere ser evaluada, pues probablemente sea consecuencia de una equivocada interpretación de los fines del experimenta de laboratorio. Un experimento de laboratorio no necesita, y no debe, constituir un intento de duplicar una situación de la vida real. Si se quisiera estudiar alga en una situación de este tipa, serla bastante tonto tomarse el trabajo de organizar un experimenta de laboratorio para reproducir dicha situación. ¿Por qué no estudiarla directamente? El experimento de laboratorio debe tratar de crear una situación en la cual se vea claramente cómo operan las variables en situaciones especialmente identificadas y definidas. El hecho de que pueda encontrarse a no tal situación en La vida real no tiene importancia. Evidentemente, nunca puede encontrarse en La vida real La situación de La mayor parte de los experimentos de laboratorio. No obstante, en el laboratorio podemos determinar can exactitud en qué medida una variable determinada afecta La conducta a actitudes en condiciones especiales o “puras”.

6.15. ¿QUE TIPO DE ESTUDIO SON LOS EXPERIMENTOS?

Debido a que analizan las relaciones entre una a varias variables independientes y una a varias dependientes y los efectos causales de las primeras sobre las segundas, son estudios explicativos (y —asimismo— abarcan correlaciones).

6.16. EMPAREJAMIENTO EN LUGAR DE ASIGNACIÓN AL AZAR

Tal y como se comentó anteriormente, otra técnica para hacer inicialmente equivalentes a los grupos es el emparejamiento. Desde luego, este método es menos preciso que la asignación al azar. Sin embargo, si se lleva a cabo con rigor, se tienen grupos grandes y se posee información que indica que los grupos no son diferentes (por ejemplo, en un experimenta sobre métodos de enseñanza, antes de iniciarla valdría la pena comparar —entre los grupos emparejados— las puntuaciones obtenidas en cursos recientes), se puede lograr un alto grado de equivalencia inicial entre grupos. Así, los diseños se representarían con una “E” de emparejamiento en lugar de la “R” (aleatorización). Par ejemplo:

E	G1	X1	01
E	G2	X2	02
E	G3	—	03

6.17. ¿QUE OTROS EXPERIMENTOS EXISTEN?: CUASIEXPERIMENTOS

Los diseños cuasiexperimentales también manipulan deliberadamente al menos una variable independiente para ver su efecto y relación con una a más variables dependientes, solamente que difieren de los experimentos “verdaderos” en el grado de seguridad a confiabilidad que pueda tenerse sobre la equivalencia inicial de los grupos. En los diseños cuasiexperimentales los sujetos no son asignados al azar a los grupos, ni emparejados; sino que dichos grupos ya estaban formados antes del experimento, son grupos intactos (La razón por la que surgen y la manera como se formaron fueron independientes a aparte del experimento). Por ejemplo, si los grupos del experimento son tres grupos escolares existentes que estaban formados con anterioridad al experimento, y cada uno de ellos constituye un grupo experimental. Veámoslo gráficamente:

Grupo A (30 estudiantes)	Grupo experimental con X ₁ .
Grupo B (26 estudiantes)	Grupo experimental con X ₂ .
Grupo C (34 estudiantes)	Grupo de control.

Otro caso sería el de que en un experimento sobre productividad en una planta, un grupo experimental fuera la cuadrilla No. 1 del turno matutino, otro grupo fuera la cuadrilla No. 2 del mismo turno, el tercer grupo fuera la cuadrilla No. 1 del turno vespertino y el grupo de control fuera la cuadrilla No. 2 del turno vespertino. Es decir, se toma a grupos constituidos. Otros ejemplos, serían utilizar grupos terapéuticos ya integrados; equipos deportivos previamente formados a grupos de habitantes de distintas zonas geográficas (que ya estaban agrupados por zona).

Problemas de los diseños cuasiexperimentales

Estos diseños se utilizan cuando no es posible asignar al azar a los sujetos a los grupos que recibirán los tratamientos experimentales. La falta de aleatorización introduce posibles problemas de validez interna y externa. Como comenta Weiss (1980, p. 89): Estos diseños tienen que luchar “con la selección como fuente posible de interpretación equivocada, la misma que con la interacción de la selección y otros factores, así como, posiblemente, con los efectos de la regresión”. Asimismo, diversos factores 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. Y dada que su validez es menor que la de los experimentos “verdaderos”, reciben el nombre de cuasiexperimentos.

Debido a los problemas potenciales de validez interna, en estos diseños el investigador debe intentar establecer la semejanza entre los grupos, esta requiere considerar las características a variables que puedan estar relacionadas con las variables estudiadas (Wiersma, 1986). Por ejemplo, si grupos intactos de trabajadores están involucrados en un

experimento sobre motivación, el turno probablemente tenga que ser introducido como una constante (grupos intactos, 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 pueda generar diferencias entre los grupos. Entre 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 en base a un criterio que pudiera afectar (v.g., formación de los salones por inteligencia) y si a los grupos en el pasado no les ha ocurrido algo que pudiera influir 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 ser identificadas con claridad, la equivalencia de los grupos tiene que discutirse y la posibilidad de generalizar los resultados así como la representatividad, deberán argumentarse sobre una base lógica (Wiersma, 1986).

Los cuasiexperimentos difieren de los experimentos “verdaderos” 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 no fueran posibles los cuasiexperimentos ya hubieran sido desechados como diseños de investigación. Más bien quiere decir que en algunos casos, los grupos pueden no ser equiparables; y el investigador debe analizar si los grupos son o no equiparables, en esta última situación el investigador debe declinar hacer la investigación con fines explicativos y limitarse a propósitos descriptivos y/o correlacionales.

Tipos de diseños cuasiexperimentales

Con excepción de la diferencia que acabamos de mencionar, los cuasiexperimentos son muy parecidos a los experimentos “verdaderos”. Por lo tanto, podemos decir que hay casi tantos diseños cuasiexperimentales como experimentales “verdaderos” solamente que no hay asignación al azar o 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 ella que nos limitaremos a ver sólo algunos de los diseños cuasiexperimentales (el resto pueden ser deducidos de sus correspondientes diseños experimentales “verdaderos”, quitándoles la “R” de asignación al azar) y serán comentados brevemente porque —insistimos— las

comparaciones, interpretaciones y análisis son iguales.³³ Los autores consideramos que sería “pecar” de repetitivos, el volver a explicar dichas comparaciones, interpretaciones y análisis.

1. Diseño con postprueba únicamente y grupos intactos

Este primer diseño utiliza a dos grupos, uno recibe el tratamiento experimental y el otro no. Los grupos son comparados en la postprueba para analizar si el tratamiento experimental tuvo un efecto sobre la variable dependiente (01 con 02). El diseño puede diagramarse del siguiente modo:

G1	X	01
G2	—	02

No hay asignación al azar ni emparejamiento.

Obsérvese que si los grupos no son equiparables entre sí, diferencias en las postpruebas de ambos grupos pueden ser atribuidas a la variable independiente pero también

³³ 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 “verdadero” correspondiente, recordando que la diferencia es que en el cuasiexperimental no hay aleatorización y los grupos pueden estar no correlacionados. a otras razones diferentes, y lo peor es que el investigador puede no darse cuenta de ella. Por ejemplo, supongamos que se lleva a cabo un cuasiexperimento para analizar el efecto de la retroalimentación que los médicos sana sus pacientes (respecta a su conducta en el tratamiento prescrita) sobre la obediencia a apego al tratamiento. Se podría partir de la siguiente hipótesis: “Los pacientes que reciban mayor retroalimentación de parte de sus médicos acerca de cómo se están comportando 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 (por ejemplo, les pidan información sobre cómo se han comportado como pacientes y les comenten dicho comportamiento, y hagan sugerencias y recomendaciones al respecto), propiciarán pacientes que en la sucesiva se apeguen más al tratamiento. Cuestión de motivar al paciente. Entonces, el investigador toma dos grupos de pacientes. Un grupo recibe retroalimentación sobre su conducta en el tratamiento prescrito y el otro grupo no. Posteriormente se evalúa que tanta se apega cada grupo —en la sucesiva— al tratamiento. Supongamos que obtenemos el siguiente resultado: 01 > 02 (el grupo experimental se apega más al tratamiento); entonces deducimos que la hipótesis fue confirmada. Pero para poder deducir lo anterior, debemos analizar muy cuidadosamente que sea posible comparar a los grupos.

Imaginemos que el grupo experimental estaba formado por pacientes que asisten a un hospital donde frecuentemente se dan pláticas motivadoras para que los pacientes 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. ¿Los resultados a qué se podrían atribuir con certeza?, ¿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 podemos tener certeza de cuál fue la causa a que tanto contribuyeron los diversos factores involucrados. Hay problema de validez interna.

También podría ser que el grupo experimental estuviera compuesto por pacientes que, desde antes del experimento, tuviera una motivación elevada para apegarse a tratamientos médicos; o podrían actuar otros factores que provocaran 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, médico que los atiende, mismas instrucciones y lugar, equivalencia como grupos en sexo, 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, puede suceder que los más enfermos se apeguen más al tratamiento). El criterio de los experimentos “verdaderos” en relación a mantener la igualdad de los grupos (salvo la manipulación de la variable independiente), se aplica igualmente a los cuasiexperimentos.

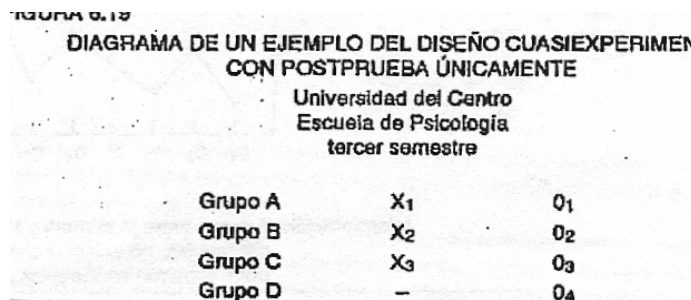
El diseño puede extenderse para incluir más de dos grupos. Teniendo así diferentes tratamientos experimentales a niveles de manipulación. Su formato general sería:

G_1	X_1	O_1
G_2	X_2	O_2
G_3	X_3	O_3
.	.	.
.	.	.
.	.	.
G_k	X_k	O_k
G_{k+1}	-	O_{k+1}

El último grupo es de control.

Un ejemplo de este diseño, sería tomar cuatro grupos escolares de un mismo semestre y carrera en una universidad, como grupos del cuasiexperimento. Veámoslo esquemáticamente en la figura 6.19.

FIGURA 6.19



Recuérdese que los grupos son intactos, no se crean, 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 preprueba-postprueba y grupos intactos (uno de ellos de control)*

Este diseño es similar al de can postprueba únicamente y grupos intactas, solamente que a las 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 sencilla serla el siguiente:

G_1	O_1	X	O_2
G_2	O_3	—	O_4

Aunque puede extenderse a más de dos grupos (niveles de manipulación de la variable independiente), La cual se podría esquematizar así:

G_1	O_1	X_1	O_2
G_2	O_3	X_2	O_4
G_3	O_5	X_3	O_6
.	.	.	.
.	.	.	.
.	.	.	.
G_K	O_{2K-1}	X_K	O_{2K}
G_{K+1}	O_{2K+1}	-	O_{2K+2}

Las posibles comparaciones entre las mediciones de La variable dependiente y las interpretaciones son las mismas que en el diseñ a experimental de preprueba-postprueba con grupo de control solamente que en este segundo diseñ a cuasiexperimental los grupos son intactos y en La interpretación de resultados debemos tomarlo en cuenta. Recuérdese toda La que se ha venido diciendo de La posible no equivalencia de los grupos. Este aspecto se aplica a todos los diseñ as cuasiexperimentales.

3. *Diseños cuasiexperimentales de series cronológicas*

En ocasiones el investigador puede pretender analizar efectos al mediano y largo plazo o efectos de administrar varias veces el tratamiento experimental y no cuenta con la posibilidad de asignar al azar a los sujetos a los grupos del experimento. En este caso, pueden utilizarse los diseños cuasiexperimentales salvo que los grupos son intactas. En ambos tipos de diseños 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 literatura clásica sobre experimentos (ver Campbell y Stanley, 1966), se reconoce coma cuasiexperimento a un diseño que no tiene grupo de control. Bien, hablemos brevemente de estos diseños.

Series cronológicas de un solo grupo

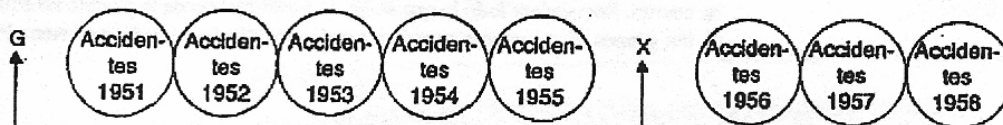
A un único grupo se le administran varias prepruebas, después se le aplica el tratamiento experimental y finalmente varias postpruebas. El diseño podría diagramarse así:

G 01 02 03 X 04 05 06

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

Un ejemplo muy difundido de este diseño lo constituyó la evaluación de un programa que tenía por objeto disminuir La velocidad en carretera del estado de Connecticut (Campbell, 1975). Los investigadores recolectaron informes y datos de accidentes de tránsito correspondientes a varios años anteriores y posteriores 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, 1980). Entonces, fue necesario comparar las estadísticas de Connecticut con las de otros cuatro estados vecinas 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 del número de accidentes. Las comparaciones dieron pie para concluir que el programa había producido efectos (Campbell, 1975; Glass, 1968).

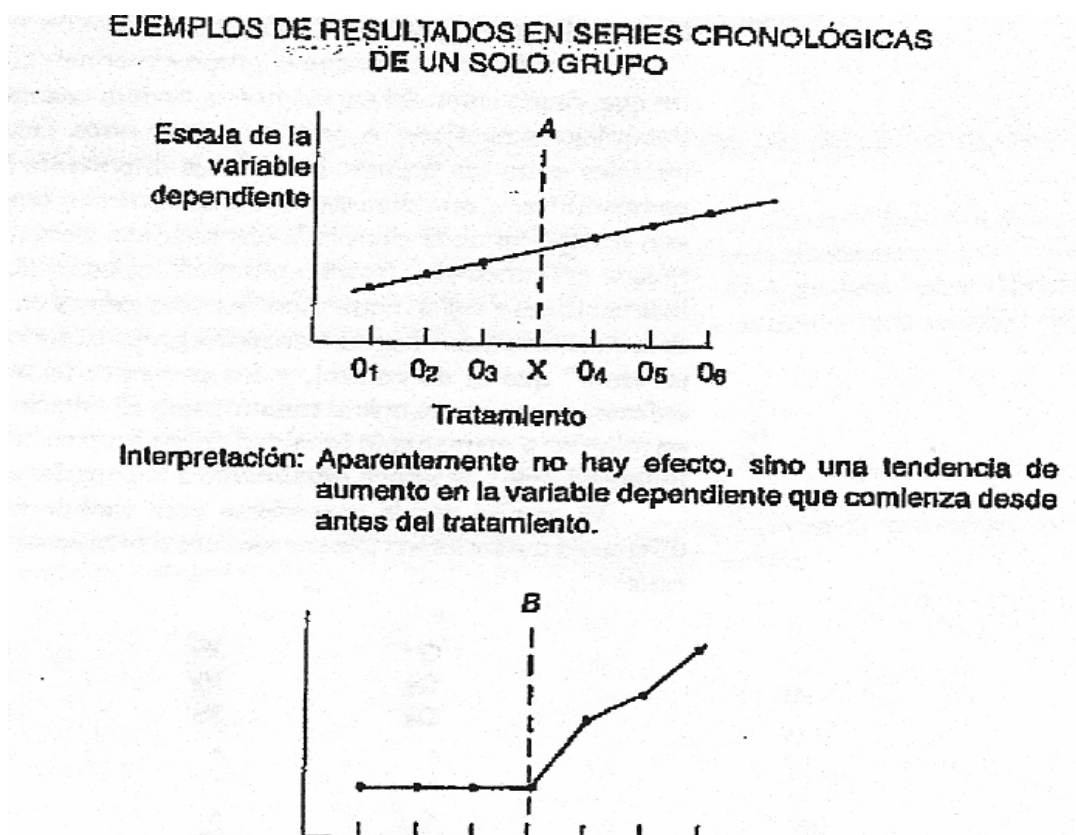
Esta investigación cuasiexperimental en su primera etapa utilizó las series cronológicas de un solo grupo:



Otro ejemplo de este diseño sería el de 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 pueden producir diversos patrones de resultados. A manera de ejemplo podríamos tener los siguientes patrones (algunos de los cuales fueron expuestas en las series cronológicas experimentales): Ver Figura 6.20.

FIGURA 6.20



En las series cronológicas de un único grupo debe tenerse muy en cuenta que no se tiene punto de comparación (grupo de control), por lo tanto la interpretación del patrón en la variable dependiente (a patrones de las variables dependientes) debe ser muy cuidadosa, habrá de analizarse si no han actuado o interactuado otros posibles causas además del tratamiento experimental a 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 son asignados al azar a los grupos, y en las cuasiexperimentales tenemos grupos intactos. Por lo tanto podríamos tener las mismas variaciones³⁴, las que se muestran en la figura 6.21.

FIGURA 6.21
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

Series cronológicas cuasiexperimentales con repetición del estímulo

Estas series también son similares a sus correspondientes experimentales, pero con grupos intactas. 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			

³⁴ El lector debe revisar los apartados relativos a las series cronológicas experimentales antes de leer este apartado. De hecho, podría notar que los diseños son los mismos, salvo que en los diagramas de las series experimentales aparece el símbolo “R” de aleatorización o 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 explicación, porque lo que puede decirse es lo mismo que en las series cronológicas experimentales. Desde luego, debe recalcar — una vez más — que en las series cuasiexperimentales los grupos son intactos y debe observarse que los grupos sean equiparables.

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 podríamos tener diagramas como éstos:

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

6.18. PASOS AL REALIZAR UN EXPERIMENTO O CUASIEXPERIMENTO

A continuación mencionamos los principales pasos que suelen realizarse en el desarrollo de un experimento a cuasiexperimento.

Paso 1: Decidir cuantas variables independientes y dependientes deberán ser incluidas en el experimento a cuasiexperimento. No necesariamente el mejor experimento es aquel que incluye el mayor número de variables, deben incluirse las variables que sean necesarias para probar las hipótesis, alcanzar los objetivos y responder las preguntas de investigación.

Paso 2: Elegir los niveles de manipulación de la(s) variable(s) independiente(s) y traducirlos en tratamientos experimentales. Este paso requiere que un concepto teórico sea convertido en una serie de operaciones que habrán de realizarse para administrar una a varios tratamientos experimentales.

Paso 3: Desarrollar el instrumento a instrumentos para medir la(s) variable(s) dependiente(s).

Paso 4: Seleccionar una muestra de personas para el experimento (idealmente representativa de la población). Véase el capítulo: “Selección de La muestra”.

Paso 5: Reclutar a los sujetos del experimento a cuasiexperimento. Esto implica tener contacto con ellos, darles las explicaciones necesarias e indicarles el lugar, día, hora y persona con quien deben presentarse. Siempre es conveniente darles el máximo de facilidades para que acudan al experimento (si se les puede brindar transporte en caso de que sea necesario, proporcionarles un mapa con las indicaciones precisas, etcétera). También hay que darles cartas (a ellos o a alguna institución a la que pertenezcan y que facilite su participación en el experimento —por ejemplo, en escuelas a los directivos, maestros y padres de familia—), llamarles por teléfono el día anterior a la realización del experimento para recordarles su participación en éste.

Los sujetos deben encontrar motivante su participación en el experimento. Por lo tanto, resulta muy conveniente darles algún regalo atractivo (a veces simbólico). Por ejemplo, a amas de casa una canasta de productos básicos, a ejecutivos una canasta pequeña con dos o tres artículos, a estudiantes créditos escolares, etcétera; y expedirles una carta de agradecimiento. A veces —por ejemplo— puede resultar adecuado que quien trate con los participantes sean personas que les resulten atractivas o atractivos (si son hombres, un grupo de muchachas bonitas puede ser muy efectivo para manejar a los sujetos en el experimento).

Paso 6: Seleccionar el diseño experimental o cuasiexperimental apropiado para nuestra(s) hipótesis, objetivos y preguntas de investigación.

Paso 7: Planear cómo vamos a manejar a los sujetos que participen en el experimento. Es decir, elaborar una ruta crítica de qué van a hacer los sujetos desde que llegan al lugar del

experimento hasta que se retiran (paso a paso).

Paso 8: En el caso de experimentos “verdaderos”, dividirlos al azar emparejarlos; y en el caso de cuasiexperimentos analizar cuidadosamente las propiedades de los grupos intactos.

Paso 9: Aplicar las prepruebas (cuando las haya), los tratamientos respectivos (cuando no se trate de grupos de control) y las postpruebas.

Asimismo, resulta conveniente tomar nota del desarrollo del experimento, llevar una bitácora minuciosa de todo lo ocurrido a lo largo de éste.

Ello nos ayudará a analizar la posible influencia de variables extrañas que generan diferencias entre los grupos y será un material invaluable para la interpretación de los resultados.

RESUMEN

1. En su acepción más general, un experimento consiste en aplicar un estímulo a un individuo o grupo de individuos y ver el efecto de ese estímulo en alguna(s) variable(s) del comportamiento de éstos. Esta observación se puede realizar en condiciones de mayor o menor control. El máximo control se alcanza en los “experimentos verdaderos”.
2. Deducimos que un estímulo afectó cuando observamos diferencias —en las variables que supuestamente serían las afectadas— entre un grupo al que se le administró dicho estímulo y un grupo al que no se le administró, siendo ambos iguales en todo excepto en esto último.
3. Para lograr el control o la validez interna los grupos que se comparen deben ser iguales en todo, menos en el hecho de que a un grupo se le administró el estímulo y a otro no. A veces graduamos la cantidad del estímulo que se administra, es decir, a distintos grupos (semejantes) les administramos diferentes grados del estímulo para observar si provocan efectos distintos.
4. La asignación al azar es —normalmente— el método preferible para lograr que los grupos del experimento sean comparables (semejantes).
5. Hay nueve fuentes de invalidación interna:
 - a. Historia.
 - b. Maduración.
 - c. Inestabilidad.
 - d. Administración de pruebas.
 - e. Instrumentación.
 - f. Regresión.

- g. Selección.
- h. Mortalidad.
- i. Interacción entre selección y maduración.
- 6. Los experimentos que hacen equivalentes a los grupos y que mantienen esta equivalencia durante el desarrollo de aquellos, controlan las fuentes de invalidación interna.
- 7. Lograr la validez interna es el objetivo metodológico y principal de todo experimento. Una vez que se consigue es ideal alcanzar validez externa (posibilidad de generalizarlos resultados a la población, otros experimentos y situaciones no experimentales).
- 8. Hay dos contextos en donde pueden realizarse los experimentos: el laboratorio y el campo.
- 9. En los cuasiexperimentos no se asignan al azar los sujetos a los grupos experimentales, sino que se trabaja con grupos intactos.
- 10. Los cuasiexperimentos alcanzan validez interna en la medida en que demuestran la equivalencia inicial de los grupos participantes y la equivalencia en el proceso de experimentación.
- 11. Los experimentos “verdaderos” constituyen estudios explicativos, los preexperimentos básicamente son estudios exploratorios y descriptivos; los cuasiexperimentos son —fundamentalmente- correlacionales aunque pueden llegar a ser explicativos.

CONCEPTOS BÁSICOS

Experimento
Cuasiexperimento
Preexperimento
Diseño experimental
Variable experimental
Variable independiente
Variable dependiente
Estímulo o tratamiento experimental/manipulación de la variable independiente
Sujetos del experimento
Control experimental
Validez interna
Influencia de variables extrañas
Validez externa
Diseño con postprueba únicamente y grupo de control
Diseño con preprueba-postprueba y grupo de control
Diseño de cuatro grupos de Solomon

Diseños experimentales de series cronológicas múltiples
Diseños experimentales de series cronológicas con repetición del estímulo
Diseños experimentales con tratamientos múltiples
Diseños factoriales
Experimento de campo
Experimento de laboratorio
Grupos intactos
Diseño con postprueba únicamente y grupos intactos
Diseño con preprueba-postprueba y grupos intactos
Diseños cuasiexperimentales de series cronológicas

EJERCICIOS

1. Seleccione una serie de variables y piense en cómo pueden manipularse en situaciones experimentales. ¿Cuántos niveles podrían incluirse para cada variable?, ¿estos niveles cómo podrían traducirse en tratamientos experimentales?, ¿se tendría un nivel de ausencia (cero) de la variable independiente?, ¿en qué consistiría éste?
2. Seleccione un experimento en alguna publicación científica (ver apéndice número 1). Analice: ¿cuál es el planteamiento del problema? (objetivos y preguntas de investigación), cuál es la hipótesis que se pretende probar a través de los resultados del experimento?, ¿cuál es la variable independiente o cuáles son las variables independientes?, ¿cuál es la variable dependiente o son las variables dependientes?, ¿cuántos grupos se incluyen en el experimento?, ¿son éstos equivalentes?, ¿cuál es el diseño que el autor o autores han elegido?, ¿se controlan las fuentes de invalidación interna?, ¿se controlan las fuentes de invalidación externa?, ¿se encontró algún efecto?
3. Un grupo de investigadores está tratando de analizar el efecto que tiene la extensión de un discurso político sobre la actitud hacia el tema tratado y al orador. La extensión del discurso es la variable independiente y tiene cuatro niveles: Dos horas, una hora y media, una hora y media hora. Las variables dependientes son la actitud hacia el tema tratado y la actitud hacia el orador (favorable-desfavorable), las cuales se medirán por pruebas que indiquen dichos niveles actitudinales. En el experimento están involucradas personas de ambos sexos, edades que fluctúan entre los 18 y los 50 años y diversas profesiones de dos distritos electorales. Existe la posibilidad de asignar al azar a los sujetos a los grupos experimentales. Desarrolle y describa dos o más diseños experimentales que puedan aplicarse al estudio, considerando a cada una de las fuentes de invalidación interna (¿puede alguna afectar los resultados del experimento?). Establezca las hipótesis que podrían establecerse para este estudio.

4. Tome un grupo de 50 personas, pídale a cada una que anoten en un trozo de papel su sexo, edad, nivel educativo, nivel mensual de ingresos, calificación que hayan obtenido en algún curso anterior y otros aspectos de interés que considere convenientes. Divida al azar a las 50 personas en dos grupos de 25 (estrictamente de manera aleatoria). Después compare cuántos hombres y mujeres hay en cada grupo; y compare los promedios de edad, años cursados, ingreso mensual, calificación en el curso elegido, etcétera, de ambos grupos. Observará que los grupos son bastante parecidos entre sí (equivalentes), las variaciones serán mínimas. El ejercicio sirve para demostrar que la aleatorización funciona.

5. Desarrolle un diseño cuasiexperimental de series cronológicas con al menos tres grupos. Describa cada grupo y tratamiento, así como el número de mediciones que se incluirán y esquematícelo de acuerdo con la simbología vista. Además establezca la hipótesis que le corresponde.

6. Considere el siguiente diseño:

R	G1	01	X1	02
R	G2	03	X2	04
R	G3	05	—	06

¿Qué podría concluirse de las siguientes comparaciones y resultados? (los signos de “igual” significan que las mediciones no difieren en sus resultados, los signos de “no igual” significan que las mediciones difieren sustancial o significativamente entre sí. Considérense sólo los resultados que se presentan y tómese en cuenta cada conjunto de resultados por separado —independientemente-).

R	G_1	O_1	X_1	O_2
R	G_2	O_3	X_2	O_4
R	G_3	O_5	-	O_6

7. Diseñe otras situaciones de interpretación y practique.

8. La investigación que ha desarrollado a lo largo de los ejercicios del libro, ¿corresponde a un experimento? Responda y explique.

BIBLIOGRAFÍA

CAMPBELL, Di (1975). Reforms as experiments, en EL. Struening y M. Guttentag SUGERIDA (Eds.): ‘Handbook of Evaluation Research’. Beverly Hills, CA: Sage Publications,

pp. 71-100. Reimpresión revisada por el autor de *American Psychologist*, vol. 24, No. 4 (abril, 1969).

CAMPBELL, DI. y STANLEY, J.C. (1966) *Experimental and quasi-experimental designs for research*. Chicago, Ill.: Rand McNally & Company. La primera edición en español (1973) fue publicada por Editorial Amorrortu.

CHRISTENSEN, L.B. (1980). *Experimental methodology*. Boston, Mass.: Allyn and Bacon, Inc. Segunda edición.

INSKO, Ch.A. y SCHOPLER, J. (1980). *Psicología Social Experimental*. México, D.F.:

Editorial Trillas.

MAIHESON, D.W.; BRUCE, R.L. y BEAUCHAMP, K.L. (1985). *Psicología Experimental: Diseños y análisis de investigación*. México, D.F.: Compañía Editorial Continental.

NUNNALLY, J.C. (1975). The study of change in evaluation research: Principles concerning measurement, experimental design, and analysis, en E.L. Struening y M. Guttentag (Eds), op. cit.; pp.101-137.

SOLOMON, RL. (1949). An extension of control-group design. *Psychological Bulletin*, No. 46, pp.137-150.

WIERSMA, W. (1986). *Research methods in education*. Newton, Mass.: Allyn and Bacon, Inc. Cuarta edición.