

Diseño de investigación: significado, propósito y principios

Diseño es el plan, estructura y estrategia de una investigación cuyo objetivo es dar respuesta a ciertas preguntas y controlar la varianza. El plan es un esquema global o programa que consiste en bosquejar lo que realizará el investigador, desde la redacción de las hipótesis y sus consecuencias operacionales, hasta el análisis final de los datos. La estructura es más específica: es el bosquejo, el esquema, el paradigma de operación de las variables. (1) Al trazar diagramas donde se describen las variables, su relación y yuxtaposición, construimos esquemas estructurales que nos permiten alcanzar las finalidades operacionales de la investigación. El significado de estrategia, según la acepción que se le da en esta obra, es también más específico que el de plan, pues comprende los métodos mediante los que se reunirán y analizarán los datos. En otras palabras, indica cómo se conseguirán los objetivos de la investigación y cómo se abordarán los problemas que se hallen en ella.

Propósito del diseño de investigación

El diseño tiene dos finalidades básicas: 1) dar respuestas a las preguntas de investigación, y 2) controlar la varianza. Desde luego, el diseño por sí solo no hace esto, sino el investigador. El diseño ayuda a contestar las interrogantes y a controlar las varianzas experimentales, extrañas y de error que surgen en el problema bajo estudio. Toda actividad de investigación se propone responder a las preguntas planteadas; en consecuencia, se puede omitir este objetivo en nuestra exposición y afirmar que se persigue un propósito fundamental: controlar la varianza. Sin embargo, semejante delimitación de objetivos es peligrosa; si no se hace hincapié en las preguntas de investigación y en el empleo del diseño para obtener las respuestas, el estudio puede dege-

nerar en un ejercicio técnico muy interesante, pero estéril.

Los diseños de investigación se elaboran para facilitar la tarea del investigador, a fin de que pueda dar respuesta a las preguntas que dieron origen a la indagación, con el máximo de exactitud y economía de esfuerzo. Todo plan de investigación se concibe y ejecuta en forma deliberada y específica para recabar la evidencia empírica relacionada con el problema en cuestión. Los problemas de investigación pueden enunciarse, y de hecho se enuncian, en forma de hipótesis; en cierta etapa, se plantean de modo que puedan someterse a verificación empírica. Contamos con una amplia gama de formas de comprobación posibles, al menos en teoría se tiene el mismo número de diseños que de posibilidades. Los diseños se preparan cuidadosamente para obtener respuestas confiables y válidas a las preguntas de investigación resumidas en la hipótesis. Puede hacerse una observación e inferir que existe la relación supuesta, pero es obvio que no se puede aceptar la inferencia como un hecho. Por otra parte, también es posible efectuar centenares de observaciones y deducir, a partir de ellas, que existen las relaciones supuestas. En tal caso podemos aceptar o rechazar la validez de la inferencia. El resultado dependerá de la manera en que se realicen las observaciones y las inferencias. Un diseño debidamente planeado y ejecutado es de gran ayuda, ya que nos permite basarnos tanto en nuestras observaciones como en nuestras inferencias.

¿Cómo se logra esto por medio del diseño? En primer lugar habría que decir que el diseño de investigación establece el marco de referencia para hacer pruebas adecuadas de las relaciones existentes entre las variables. Nos dice, en cierto modo, cuáles observaciones hemos de efectuar, cómo hacerlas y la manera de analizar sus representaciones cuantitativas. En rigor, no "indica" exactamente qué hemos de hacer, sino que "nos señala" las direcciones a seguir para efectuar la observación y el análisis. Un buen diseño "nos sugiere", vaya por caso, cuántas observaciones es preciso realizar y cuáles variables son activas o atributivas. Entonces podemos manipular unas y clasificar otras. Un diseño nos indica qué tipo de análisis estadístico se debe usar. Por

último, nos señala las posibles conclusiones que se extraerán del análisis estadístico.

Un ejemplo

Se dice que los colegios y universidades estadounidenses discriminan a las mujeres tanto en la admisión de estudiantes como en la contratación de personal. (2) Supongamos que queremos verificar la primera clase de discriminación. Para ello llevamos a cabo un experimento en la forma siguiente. Se escoge al azar una muestra de 200 colegios y se envían solicitudes de admisión, con base en varios casos modelo seleccionados según una gama de habilidades comprobadas, y hacemos que todos los detalles sean iguales menos el sexo. Una mitad de las solicitudes corresponderá a los varones y la otra mitad a las mujeres. En igualdad de circunstancias, cabe esperar números aproximadamente iguales de aceptaciones y rechazos. La aceptación es, pues, la variable dependiente. Se la mide sobre una escala de tres puntos: aceptación plena, aceptación con reservas, rechazo. Designemos a los varones con  $A_1$  y a las mujeres con  $A_2$ . El paradigma del diseño aparece en la figura 17-1.

El diseño es el más sencillo que se pueda imaginar, con los requisitos mínimos de control. Los dos tratamientos serán asignados aleatoriamente a los colegios de modo que cada uno reciba una solicitud de un varón o de una mujer. Con objeto de saber si tiene valor estadístico, la diferencia entre las medias,  $M_{A_1}$  y  $M_{A_2}$ , se comprobará mediante una prueba  $t$  o  $F$ . La hipótesis esencial es:  $M_{A_1} > M_{A_2}$ , o sea que se aceptará a un número mayor de varones que de mujeres. Si no hay discriminación en las admisiones,  $M_{A_1} = M_{A_2}$  desde el punto de vista estadístico. Supongamos que una prueba  $F$  indica que difieren poco las medias, ¿se podría asegurar entonces que en general no se practica la discriminación? Si bien el diseño presentado en la figura 17-1 es satisfactorio tal como está, se podría mejorar.

Un diseño más completo

Walster y sus colegas (véase la nota 2), en su estudio sobre la discriminación que se practica en las admisiones, se valieron de otras dos variables independientes (raza y habilidad) en un diseño factorial. Hacemos a un lado la raza —pues no tuvo valor estadístico ni interactuó considerablemente con las demás variables— y nos concentramos en el sexo y en la habilidad. Si la selección de los estudiantes de nuevo ingreso hecha por un colegio se funda exclusivamente en la habilidad, no habrá discriminación

(2) La idea para el ejemplo se tomó de un experimento original e ingenioso: E. Walster, T. Cleary y M. Clifford, "The Effect of Race and Sex on College Admission," *Sociology of Education*, XLIV (1971), 237-244.

Tratamientos	
$A_1$ (Varón)	$A_2$ (Mujer)
ACEPTACION	
PUNTUACIONES	
$M_{A_1}$	$M_{A_2}$

Figura 17-1.

(a menos, claro está, que a tal procedimiento se le designe así). Añádase la variable "habilidad" al diseño de la figura 17-1; úsense tres niveles. Es decir, las solicitudes no sólo se designarán en función del sexo, sino que además se designarán en base a la habilidad, alta, mediana y baja. Así, tres de las solicitudes serían: varón, capacidad mediana; mujer, capacidad alta; mujer, capacidad baja. Ahora bien, si no hubiera diferencia significativa entre ambos sexos y en la interacción entre sexo y habilidad, se obtendría una evidencia más fuerte respecto a la ausencia de discriminación que la aportada por el diseño y la prueba estadística de la figura 17-1. Con el diseño ampliado, figura 17-2, vamos a explicar esta afirmación y a exponer varios puntos referentes al diseño de investigación.

Es un diseño factorial de tipo  $2 \times 3$ . Una variable independiente,  $A$ , es el sexo, la misma que en la figura 17-1. La segunda variable independiente,  $B$ , es la habilidad, la cual se manipula para indicar, en diversas formas, cuáles son los niveles de habilidad que poseen los estudiantes. (3) Supongamos que, en nuestra opinión, la discriminación contra las mujeres asume una forma más sutil que la exclusión manifiesta: a las de menor habilidad se las discrimina (en comparación con los varones). Se trata entonces de una hipótesis de interacción. En definitiva, este problema y el paradigma de la figura 17-2 nos servirán de base para comentar algunos elementos del diseño de investigación.

Cada problema de investigación sugiere un diseño específico. La que acabamos de mencionar es una hipótesis de interacción, y en consecuencia el diseño factorial es de manera evidente el más apropiado.

(3) Es importante no confundir los nombres de las variables. Sexo y habilidad suelen ser variables atributivas y por lo mismo no experimentales. No obstante, en este caso son manipuladas. Los históricos de cada estudiante enviados a los colegios fueron sistemáticamente ajustados para que encajasen en las seis casillas de la figura 17-2. Un caso dentro de la casilla  $A_1B_1$ , por ejemplo, sería historial de un varón de capacidad mediana. Dichos antecedentes son lo que el colegio valora al decidir la admisión o el rechazo.

(1) Por "paradigma" se entiende un modelo o un ejemplo. La palabra "modelo" es sinónimo de paradigma, pero este vocablo no contiene la connotación valorativa de aquél. Lo utilizaremos en el sentido de una estructura y modelo orientador, en particular al tratar del diseño de investigación.

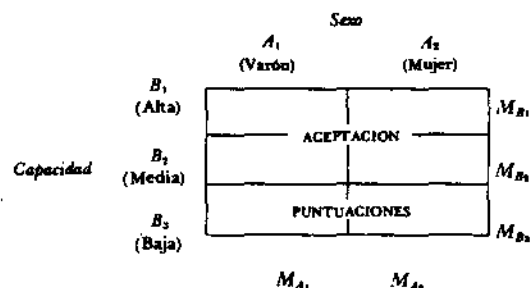


Figura 17-2.

Con la literal *A* se representa al *sexo*; con *B* la *habilidad*; *A* se divide en *A*<sub>1</sub> y *A*<sub>2</sub>, y *B* en *B*<sub>1</sub>, *B*<sub>2</sub> y *B*<sub>3</sub>.

El paradigma de la figura 17-2 sugiere varias cosas. En primer lugar, es claro que se requiere un número bastante grande de sujetos. Más exactamente se necesitan *6n* (*n* = número de sujetos en cada casilla). Si decidimos que *n* sea 10, hemos de tener 60 sujetos para el experimento. Adviértase "la sagacidad" de este diseño. Si sólo se sometieran a prueba los tratamientos y se ignorara la habilidad, únicamente necesitaríamos *2n* sujetos.

En segundo lugar, el diseño indica que los "sujetos" (colegios, en este caso) pueden ser asignados aleatoriamente a *A* y *B*, porque ambas letras representan variables experimentales. En cambio, si la habilidad fuese una variable atributiva no experimental, los sujetos podrían ser asignados aleatoriamente aun a *A*<sub>1</sub> y *A*<sub>2</sub>, pero no a *B*<sub>1</sub>, *B*<sub>2</sub> o *B*<sub>3</sub>. En tercer lugar, el diseño establece que las observaciones realizadas en los "sujetos" han de practicarse por separado. La puntuación de un colegio no ha de afectar a la de otro. Este es un requisito estadístico del diseño factorial. El simple acto de reducir un diseño a un esquema como el de la figura 17-2 prescribe en efecto las operaciones indispensables para obtener las medidas más adecuadas al análisis estadístico. Una prueba *F* se funda en la suposición de la independencia de las medidas de la variable dependiente. Si en ella la habilidad es una variable atributiva y los individuos son medidos para precisar su grado de inteligencia, por ejemplo, el requisito de independencia corre el peligro de no cumplirse por la posibilidad de que un sujeto copie las respuestas de otro, o porque los maestros inadvertida o deliberadamente les "ayuden" a contestar las preguntas y por otras razones. Los investigadores procuran impedir cosas, no sólo por motivos éticos, sino por el deseo de satisfacer los requisitos de un diseño y de desarrollar estadísticas dignas de confianza.

El cuarto punto debe estar ya muy claro para nosotros: la figura 17-2 indica un análisis factorial de

varianza, prueba *F*, medidas de asociación y, quizás, pruebas post hoc. Si antes de reunir los datos la investigación ha sido bien planeada—como en el caso de Walster y otros—, la mayor parte de los problemas estadísticos podrán ser resueltos. Además, es posible evitar algunos problemas arduos antes que surjan y hasta impedir que se presenten. En cambio, con un diseño inadecuado los problemas referentes a las pruebas estadísticas idóneas presentarán muchas dificultades. Una de las razones por las cuales en este libro se insiste en tratar concomitantemente al diseño y a los problemas estadísticos es para señalar las formas de evitar posibles dificultades. Si el diseño y el análisis estadístico se planean de manera simultánea, el trabajo analítico será sencillo y ordenado.

La última función del diseño que se expondrá aquí es la siguiente: permite bosquejar claramente las posibles conclusiones a las que puede llegar el investigador, mediante la descripción o sugerencia específica de las pruebas que pueden llevarse a cabo. Un diseño simple al azar de una variable con dos divisiones, por ejemplo, dos tratamientos, *A*<sub>1</sub> y *A*<sub>2</sub>, permite realizar únicamente una prueba estadística de la diferencia existente entre las dos estadísticas arrojadas por los datos. Las estadísticas pueden ser dos medias, dos medianas, dos intervalos, dos varianzas, dos porcentajes, etc. Normalmente sólo es posible una prueba estadística. Sin embargo, con el diseño de la figura 17-2 se pueden practicar tres: 1) entre *A*<sub>1</sub> y *A*<sub>2</sub>; 2) entre *B*<sub>1</sub>, *B*<sub>2</sub> y *B*<sub>3</sub>; 3) la interacción de *A* y *B*. En las investigaciones no todas las pruebas estadísticas tienen igual importancia. Las primordiales, por supuesto, son las que guardan relación directa con los problemas de investigación y con las hipótesis de trabajo.

En el presente caso, la hipótesis de interacción (el número 3) es la importante, puesto que la discriminación dependerá del nivel de habilidad; y es probable que los colegios la practiquen en distintos niveles. Dijimos antes que las mujeres (*A*<sub>2</sub>) serán acep-

tadas más que los varones (*A*<sub>1</sub>) en un nivel superior de capacidad (*B*<sub>1</sub>), en tanto que se les aceptará menos en un nivel inferior (*B*<sub>3</sub>). Ello se puede expresar simbólicamente en la siguiente forma:

$$H_3: A_2 > A_1 | B_1 \quad (O: A_1 = A_2 | B_1) \\ A_1 = A_2 | B_2 \quad (?) \\ A_1 > A_2 | B_3$$

Tratamientos de la condición B <sub>1</sub>	
A <sub>1</sub>	A <sub>2</sub>

Tratamientos de la condición B <sub>2</sub>	
A <sub>1</sub>	A <sub>2</sub>

M <sub>B<sub>3</sub>A<sub>1</sub></sub>	M <sub>B<sub>3</sub>A<sub>2</sub></sub>
---	---

M <sub>A<sub>1</sub></sub>	M <sub>A<sub>2</sub></sub>
----------------------------	----------------------------

Figura 17-3.

lo cual simplemente corrobora lo que se afirmó antes, con la excepción de que la segunda hilera no es segura. En el simbolismo, el signo vertical " | " significa "en la condición" y así la primera hilera se lee: "*A*<sub>2</sub> es mayor que *A*<sub>1</sub> en la condición *B*<sub>1</sub>".

A estas alturas es evidente que el diseño no es estático. Su conocimiento nos ayuda a planear, a realizar mejores investigaciones, y además nos sugiere la comprobación de la hipótesis y tal vez lo más importante: quizá lleguemos a darnos cuenta que el diseño de un estudio no corresponde por completo a lo que se espera de él. ¿Qué queremos decir con esta afirmación tan peculiar?

Supongamos que enunciarnos la hipótesis de interacción en la forma antes descrita, sin saber siquiera que existe el diseño factorial. Elaboramos un diseño que consta en realidad de dos experimentos. En uno comparamos a *A*<sub>1</sub> con *A*<sub>2</sub> en la condición *B*<sub>1</sub>. En el segundo comparamos a *A*<sub>1</sub> con *A*<sub>2</sub> en la condición de *B*<sub>3</sub>. El paradigma será semejante al de la figura 17-3. (Para simplificar las cosas estamos utilizando únicamente dos niveles de *B*, *B*<sub>1</sub> y *B*<sub>3</sub>, con lo cual el diseño se reduce a 2 × 2.)

Se debe advertir que con este diseño no es posible realizar ninguna prueba adecuada de la hipótesis. Es verdad que *A*<sub>1</sub> puede compararse con *A*<sub>2</sub> en las condiciones *B*<sub>1</sub> y *B*<sub>3</sub>. Pero no es posible conocer, con claridad y sin ambigüedades, si hay una interacción apreciable entre *A* y *B*. Aun si según lo previsto *M*<sub>B<sub>1</sub>A<sub>1</sub></sub> > *M*<sub>B<sub>3</sub>A<sub>1</sub></sub>, el diseño no ofrece una confirmación segura de que exista la interacción, pues no podemos obtener información alguna sobre la diferencia entre *B*<sub>1</sub> y *B*<sub>3</sub>. Esa información es indispensable para comprobar la hipótesis de interacción. Si las condiciones estadísticas estuvieran indicadas en la expresión anterior, habría razones para creer en la supuesta veracidad de dicha hipótesis. Sin embargo, este tipo de evidencia no es suficiente, en especial si sabemos que es posible mejorarla.

Supongamos que en la figura 17-3 las medias de las casillas fueran, de izquierda a derecha, 30, 30; 40 y 30. Parecería que el resultado apoya la hipótesis de interacción, pues hay una apreciable diferencia entre *A*<sub>1</sub> y *A*<sub>2</sub> en el nivel de *B*<sub>3</sub>, no así en el nivel *B*<sub>1</sub>. Pero no podríamos estar absolutamente seguros, aun cuando la diferencia entre *A*<sub>1</sub> y *A*<sub>2</sub> fuese estadísticamente significativa. En la figura 17-4 se muestra

cómo se representaría esto si se utilizase un diseño factorial. (Las cifras anotadas en las casillas y en los márgenes son medias.) Supongamos que los efectos principales, *A*<sub>1</sub> y *A*<sub>2</sub>; *B*<sub>1</sub> y *B*<sub>3</sub>, fuesen significativos; aun así, es posible que la interacción no fuese importante. A no ser que la hipótesis de interacción haya sido específicamente puesta a prueba, su evidencia será probable porque falta la estructura de una prueba estadística planeada que sólo se obtiene por medio del diseño factorial. Es claro que el conocimiento del diseño podría haber mejorado este experimento.

### El diseño de investigación como control de varianza

La principal función técnica del diseño es controlar la *varianza*. Un diseño es, por decirlo así, un conjunto de instrucciones para que el investigador reúna y analice sus datos en determinada forma. Es, pues, un mecanismo regulador. El principio estadístico en que descansa, como afirmamos antes, es éste: *maximizar la varianza sistemática, regular la varianza sistemática extraña y minimizar la varianza de error*. Dicho de otra manera, se debe *controlar la varianza*.

Conforme a ese principio, al construir un diseño eficiente, el investigador procura: 1) maximizar la varianza de la variable o variables de su hipótesis fundamental; 2) controlar la varianza de variables extrañas o "indeseables" que pueden influir en los resultados experimentales y que no le interesan en ese momento, y 3) minimizar la varianza de error o varianza aleatoria, incluidos los llamados errores de medición. Examinemos un ejemplo.

#### Ejemplo

Un investigador de la educación decide demostrar la hipótesis de que la enseñanza de la lectura se logra mejor por el método "oportunistico" que por el método sistemático. (4) El método sistemático, que

(4) La idea de este ejemplo fue tomada de A. Gates, M. Batchelder y J. Betzner, "A Modern Systematic versus an Opportunistic Method of Teaching," *Teachers College Record*, XXVII (1926), 679-700. Las nociones del diseño empleadas en el ejemplo no son las de los autores, quienes recurrieron al método de igualamiento.

	$A_1$	$A_2$	
$B_1$	30	30	30
$B_2$	40	30	35
	35	30	

Figura 17-4.

se denotará por  $A_1$ , se vale de un programa de estudio que se organiza y se orienta fundamentalmente hacia una asignatura. El método *oportunistico*, que se identificará por  $A_2$ , explora las inclinaciones e intereses de los alumnos. El maestro aprovecha las oportunidades que se le presentan para introducir el conocimiento de la lectura, pues supone que el estudio estará motivado por el propio interés de los niños. El investigador conoce perfectamente que otras posibles variables independientes influirán en el resultado de la lectura: inteligencia, sexo, antecedentes de clase social, experiencia previa con materiales verbales, entre otras. Además, tiene motivos para pensar que los métodos van a funcionar en forma distinta con diferentes clases de alumnos. Quizás funcionen de una manera con los niños con diferentes grados de inteligencia y de otra manera con los que provienen de distintos ambientes familiares. Los que proceden de hogares en los que impera una disciplina estricta pueden responder mejor a un método que a otro, y se puede afirmar lo mismo de quienes provienen de hogares relativamente tolerantes.

¿Qué clase de diseño debe escoger el investigador? Para contestar esta pregunta es importante reunir y clasificar las variables y conocer con claridad cuáles preguntas desea él responder. He aquí las variables:

Métodos	Tipo de hogar
sistemático, $A_1$	restrictivo, $B_1$
oportunistico, $A_2$	liberal, $B_2$

Podría incluir asimismo la *inteligencia* como variable independiente, pero prefiere no hacerlo. En cambio, decide que la asignación al azar se ocupará de la inteligencia y de otras variables independientes capaces de influir. La medida de su variable dependiente la aporta una prueba estandarizada de lectura que se aplicará a los cuatro meses de terminado el experimento.

El problema parece exigir un diseño factorial. Tal opción obedece a dos causas: la primera es que hay dos variables independientes; la segunda es que el investigador tiene en mente una hipótesis de interacción, aunque no la exprese con estas palabras, pues

cree que los métodos no funcionarán igual con distintas clases de niños. En la figura 17-5 se muestra la estructura del diseño.

Obsérvese que las medias marginales y de las casillas fueron clasificadas debidamente. Nótese asimismo que hay una *variable activa*, varios métodos y una *variable atributiva* (el tipo de hogar). (5) El experimentador, claro está, no puede usar la aleatorización con la variable atributiva. Lo único que puede hacer es clasificar a los sujetos según el tipo de hogares a los que pertenecen, restrictivos y liberales, y asignarlos a  $B_1$  y  $B_2$ , respectivamente. No obstante, los puede asignar aleatoriamente a  $A_1$  y  $A_2$ , que representan a los dos métodos. Esto lo hace en dos etapas: 1) asigna aleatoriamente a los niños de  $B_1$ , o sea los que proceden de hogares restrictivos a  $A_1$  y  $A_2$ , y 2) asigna los niños de  $B_2$ , o sea los que provienen de hogares liberales, a  $A_1$  y  $A_2$ . Esto le permite suponer que antes de iniciar el experimento, los niños en  $A_1$  son aproximadamente iguales a los de  $A_2$  en todas las características posibles.

Por ahora lo que nos interesa son las distintas funciones de la varianza en el diseño de investigación y el principio de varianza. Antes de adentrarnos en la materia, a este último lo designaremos como principio "maxmincon", para una fácil referencia. Su origen es evidente: *maximizar* la varianza sistemática que se estudia; *controlar* la varianza sistemática extraña; y *minimizar* la varianza de error (con dos de las sílabas interpuestas por razones de eufonía).

Antes de abordar la aplicación del referido principio en el ejemplo presente, conviene aclarar un punto importante. Cada vez que hablemos de la varianza, debemos estar seguros de saber a *cuál* nos referimos. Hablamos de la varianza de los métodos, de la inteligencia, del sexo, del tipo de hogar, etc. Parece que se estuviera hablando sobre la varianza de la variable independiente. Lo cual es cierto y no. Al final de un experimento siempre nos referimos a la *varianza de la variable dependiente*, o sea la *varianza de las medidas de la variable dependiente*. (6) La expresión "varianza de la variable independiente" se debe al hecho de que mediante la manipulación y el control de las variables independientes *influyamos*, posiblemente, en la varianza de la variable dependiente. Dicho en forma un tanto inexacta, "hacemos" que las medidas de la variable depen-

(5) En el capítulo 3 se dijo que la *variable activa* es una variable experimental o manipulada, y que la *variable atributiva* es una variable medida o que designa alguna característica; por ejemplo, inteligencia, clase social y ocupación (personas); coherencia, productividad y atmósfera restrictiva-liberal (organizaciones, grupos y otros).

(6) No sucede así en los llamados estudios correlacionales, en los cuales cuando decimos "la varianza de la variable independiente" queremos decir exactamente eso. Al correlacionar dos variables estudiamos "directamente" las varianzas de las variables independientes y dependientes.

		Métodos		
		$A_1$ (sistemático)	$A_2$ (oportunistico)	
Tipo de hogar	$B_1$ (restrictivo)	$M_{A_1B_1}$	$M_{A_2B_1}$	$M_{B_1}$
	$B_2$ (liberal)	$M_{A_1B_2}$	$M_{A_2B_2}$	$M_{B_2}$
		PUNTAJES DE LECTURA		
		$M_{A_1}$	$M_{A_2}$	

Figura 17-5.

diente se comporten y varíen como un resultado supuesto de nuestra manipulación y de nuestro control de las variables independientes. En un experimento se analizan las medidas de la variable dependiente. Después inferimos del análisis que las varianzas, presentes en la varianza total de las medidas de la variable dependiente, son causadas por la manipulación y control de las variables independientes y por el error. Dicho esto, retornamos a nuestro principio.

### Maximización de la varianza experimental

La preocupación más obvia del experimentador, si bien no necesariamente la más importante, es maximizar lo que llamaremos *varianza experimental*, término que se ha introducido para facilitar las exposiciones posteriores y que, en general, simplemente designa la varianza de la variable dependiente influida por la variable o variables independientes de la hipótesis substantiva. En este caso particular, la varianza experimental es la varianza de la variable dependiente supuestamente ocasionada por los métodos,  $A_1$  y  $A_2$ , y por los tipos de hogar,  $B_1$  y  $B_2$ . Aunque por varianza experimental puede entenderse únicamente la debida a una variable manipulada o *activa*, como los métodos, también consideraremos variables experimentales a las variables *atributivas*, v. gr.: inteligencia, sexo, y, en este caso el tipo de hogar. Una de las principales tareas del experimentador consiste en maximizar dicha varianza; debe separar lo más que pueda los métodos  $A_1$  y  $A_2$  (y  $A_3$ ,  $A_4$ , etc.), si se encuentran en el diseño y hacer que difieran al máximo.

Si la variable independiente no varía substancialmente, resulta en extremo difícil separar su efecto de la varianza total de la variable dependiente, gran parte de la cual se debe a la casualidad. Hay que darle a la varianza de una relación la oportunidad de

manifestarse por sí misma y de separarse de la varianza total, que es una combinación de las varianzas derivadas de numerosas causas y del azar. Conforme a este subprincipio del principio maxmincon escribimos la siguiente regla de investigación: *diseñe, planee y realice la investigación de manera que las condiciones experimentales sean lo más distintas que se pueda.* (7)

En el ejemplo que venimos comentando, el subprincipio significa que el investigador se esforzará para hacer que difieran al máximo  $A_1$  y  $A_2$ , en este ejemplo, los métodos sistemático y "oportunistico". A continuación, es necesario clasificar los tipos de hogares en restrictivos y liberales,  $B_1$  y  $B_2$ , de manera que sean tan diferentes como sea posible en la dimensión restrictivo-liberal; el problema esencialmente es de medición. En un experimento, el investigador obra como el titiritero que mueve a placer las marionetas, o sea la variable independiente. En la mano derecha sostiene los hilos de las marionetas  $A_1$  y  $A_2$  y en la izquierda, los hilos de las marionetas  $B_1$  y  $B_2$ . (Damos por descontado que una mano no influye en la otra, es decir, que cada mano es independiente.) Hace danzar por separado a las marionetas  $A_1$  y  $A_2$ , y a las marionetas  $B_1$  y  $B_2$ . Luego contempla a su auditorio (la variable dependiente) para ver y medir el efecto de sus manipulaciones. Si logra que dancen aparte  $A_1$  y  $A_2$  y si existe una relación entre  $A$  y la variable dependiente, la reacción de los expectadores (por ejemplo, si la separación de  $A_1$  y  $A_2$  produce hilaridad), será la risa. Quizá hasta observe que sólo se ríen cuando  $A_1$  y  $A_2$  dancan aparte, mientras que  $B_1$  y  $B_2$  dancan por su lado (de nuevo una interacción).

(7) Desde luego se dan excepciones de este subprincipio, aunque probablemente son escasas. Quizá un investigador desee estudiar los efectos de gradaciones pequeñas, digamos, de los incentivos motivacionales en el aprendizaje de alguna asignatura. En este caso, no hará sus condiciones experimentales lo más distintas posible, pero tendrá que hacerlas variar un poco o no habrá ninguna varianza apreciable en la variable dependiente.

### Control de variables extrañas

El procedimiento mediante el cual las influencias de las variables independientes ajenas a las finalidades del estudio se reducen al mínimo, se nulifican o se aíslan, recibe el nombre de control de las variables extrañas. Dicho en otras palabras, la varianza de estas variables se reduce realmente a cero o casi cero o, lo que es lo mismo, se separa de la varianza de las demás variables independientes.

Disponemos de tres maneras para controlar las variables extrañas. La primera es la más fácil cuando es posible emplearla: eliminar la variable como tal. Si nos preocupa la inteligencia por ser un factor que quizá repercuta en el aprovechamiento escolar, su efecto en la variable dependiente quedará prácticamente suprimido si nos valemos de sujetos de un solo nivel de inteligencia, digamos, que tengan un cociente intelectual (IQ) dentro de un intervalo de 90 a 110. Si lo que se estudia es el aprovechamiento y si la raza es un factor capaz de influir en la varianza del mismo, se le puede eliminar utilizando exclusivamente a sujetos de una raza. Entonces, el subprincipio es: *para suprimir el efecto de una variable independiente capaz de influir en la variable dependiente, se pueden escoger sujetos con la mayor homogeneidad posible en esa variable independiente.*

Esta manera de controlar la varianza indeseable o extraña es sumamente eficaz. Si escogemos únicamente un sexo para el experimento, estaremos seguros de que el sexo no puede ser una variable independiente con influencia. Pero ello va en detrimento de la generalización; por ejemplo, no se podrá afirmar sobre las mujeres si en el experimento sólo se incluyen varones. Si se restringe el rango de inteligencia, las conclusiones tendrán que limitarse a él. ¿Es posible que la relación, en caso de ser descubierta, sea nula o muy distinta en los niños de alta inteligencia y en los de baja inteligencia? Simplemente lo ignoramos: en este caso lo único que nos queda es hacer suposiciones o conjeturas.

El segundo procedimiento de control de la varianza extraña se basa en la aleatorización. En teoría, es el único método por el que se pueden controlar todas las posibles variables extrañas. He aquí otra forma de expresarlo: si la aleatorización se ha efectuado a conciencia podrá considerarse que los grupos experimentales son estadísticamente iguales en todos aspectos. Ello no significa, claro está, que los grupos sean iguales para todas las variables posibles. Ya hemos visto que fortuitamente pueden ser desiguales; pero con una apropiada aleatorización, la probabilidad de que sean iguales será mayor. Por tal razón, el control de la varianza extraña mediante dicho procedimiento constituye un método eficaz de control. Los demás métodos dan margen a muchas desigualdades. Si apareamos a los sujetos en inteligencia, se podría conseguir igualdad estadística en ella (al menos en los aspectos de la inteligencia que

han sido medidos), pero nos toparemos con desigualdad en otras variables independientes de influencia, entre ellas: aptitud, motivación y clase social. En seguida citamos un subprincipio que emana de esa propiedad igualadora de la aleatorización: *siempre que sea posible hacerlo, asignense aleatoriamente los sujetos a los grupos y condiciones experimentales, asimismo asignense las condiciones y otros factores a los grupos experimentales.*

El tercer modo de controlar una variable extraña consiste en introducirla dentro del diseño en forma de variable independiente. Por ejemplo, supongamos que se desea controlar el sexo en el experimento citado, y que se estima inoportuno o inconveniente eliminarlo. Se podría añadir una tercera variable independiente, el sexo. No obstante, es poco probable que se recurriera a esa forma de control, a menos que nos interesara conocer la diferencia concreta entre los sexos, en cuanto a la variable dependiente, o quisiésemos estudiar la interacción entre una o dos variables más y el sexo. Quizás se pretendiera reunir información como la que acabamos de mencionar y al mismo tiempo controlar el sexo. En tal caso sería aconsejable incorporarlo al diseño. Lo importante es que al integrar una variable en el diseño experimental se puede ejercer control sobre ella, puesto que entonces es posible extraer de la varianza total de la variable dependiente la varianza debida a dicha variable. (En el caso anterior sería la varianza entre sexos.)

Tales consideraciones conducen a otro subprincipio: *una variable extraña es controlable si se la integra en el diseño en forma de variable atributiva, con lo cual se logra su control y se obtiene más información acerca de su efecto sobre la variable dependiente y acerca de su posible interacción con otras variables independientes.*

La cuarta manera de controlar la varianza extraña es aparear a los sujetos. El principio regulador en que descansa el apareamiento es idéntico al de cualquier otra clase de control, el control de varianza. El apareamiento es semejante —y de hecho se le podría considerar un corolario— al principio según el cual la varianza de una variable extraña se controla mediante su introducción en el diseño. El principio básico consiste en dividir una variable en dos o más partes (es decir, en inteligencia alta y baja cuando se trata de un diseño factorial) y luego efectuar la asignación al azar dentro de cada nivel tal como se describió previamente. El apareamiento es un caso especial de este principio. Pero en vez de dividir a los sujetos en dos, tres o cuatro partes, se dividen en  $N/2$  partes, donde  $N$  es el número de sujetos utilizados; de ese modo se identifica y se incorpora al diseño el control de varianza.

Al utilizar el procedimiento de apareamiento se presentan varios problemas. Ante todo, la variable en que se aparean los sujetos debe guardar una rela-

ción substancial con la variable dependiente, o de lo contrario, el apareamiento será una pérdida de tiempo. Y lo que es peor, podrá conducir a conclusiones erróneas. Además, adolecerá de graves limitaciones: si tratamos de aparear una o más variables y hasta más de una, perderemos a los sujetos. Por ejemplo, si se decide igualar inteligencia, sexo y clase social, acaso tengamos éxito al aparear las primeras dos variables pero no al tratar de encontrar pares que sean suficientemente iguales en las tres variables. Agréguese una cuarta variable y el problema se volverá aún más difícil, y a menudo imposible de resolver.

Pero no permitamos que los inconvenientes nos hagan perder de vista las ventajas. Cuando se da una correlación substancial de la variable o variables igualadas con la variable dependiente ( $>.50$  ó  $.60$ ), esta técnica reduce el término de error y con ello aumenta la precisión del experimento. Si participan los mismos sujetos en distintos tratamientos experimentales —método llamado diseño de medidas repetidas o bloques asignados al azar— dispondremos de un potente control de varianza. ¿Hay acaso una forma mejor de igualar todas las variables posibles que aparear a cada sujeto consigo mismo? Por desgracia, otras consideraciones negativas suelen excluir esta posibilidad. Hay que hacer hincapié en que ningún tipo de apareamiento es capaz de substituir a la aleatorización. Una vez que los sujetos han sido igualados, es preciso asignarlos aleatoriamente a los grupos experimentales. Por medio de un procedimiento al azar, como arrojar una moneda o el uso de números impares e incluso aleatorios, los miembros de los pares se distribuyen en los grupos experimental y testigo de control. Si los mismos sujetos se someten a todos los tratamientos, el orden de éstos deberá asignarse al azar y así el control del apareamiento o repetición de medidas, se incrementa gracias al control de la "aleatorización". (8)

Un subprincipio contenido en nuestra exposición es el siguiente: *cuando una variable de apareamiento guarda correlación substancial con la variable dependiente, dicho apareamiento, como forma de control de varianza, puede ser útil y deseable. No obstante, antes de recurrir a él deben ponderarse con cuidado sus ventajas y desventajas en cada situación particular de investigación. La aleatorización completa o el análisis de covarianza podrían ser los métodos más idóneos de control de varianza.*

Otro método de control, el estadístico, se contentó con amplitud en la parte 5; pero ahora se harán dos o más observaciones con mayor detalle. Por así decirlo, los métodos estadísticos son formas de control en el sentido de que aíslan y cuantifican las varian-

zas, pero son inseparables de otras formas de control de diseños. Por ejemplo, si se utiliza el apareamiento, hay que usar una prueba estadística idónea o de lo contrario se perderá el efecto del apareamiento y con él también el control.

### Minimización de la varianza de error

*Varianza del error* es la variabilidad de las medidas causada por las fluctuaciones fortuitas cuya característica fundamental es que se compensan a sí mismas, variando en esta o aquella manera; a veces son positivas, a veces negativas, a veces ascienden y otras descienden. Los errores fortuitos tienden a equilibrarse entre sí de manera que su media sea cero, pero la varianza sistemática es, en esencia, previsible. En cambio, la varianza de error es impredecible.

Hay varias determinantes de la varianza de error, tales como los factores asociados a las diferencias individuales de los sujetos. De ordinario le damos el nombre de "varianza sistemática" a la que surge por las diferencias individuales. Pero cuando es imposible identificarla y controlarla o no se hace esto, tenemos que incluirla en la varianza de error. Esta última posee esa característica aleatoria porque algunos determinantes interactúan y tienden a cancelarse mutuamente, al menos es lo que suponemos.

Otra causa de la varianza de error es la que se relaciona con los llamados errores de medición: variación de respuestas de un ensayo a otro, conjeturas, distracción momentánea, ligera fatiga temporal y lapsos de memoria, estados emocionales transitorios de los sujetos, entre otros.

La minimización de la varianza de error tiene dos aspectos principales: 1) reducción de los errores de medición mediante el control de las condiciones; 2) incremento en la confianza de las medidas. Cuanto menos controladas estén las condiciones de un experimento, mayor será la acción de los diversos determinantes de la varianza de error. Es uno de los motivos por los que deben establecerse situaciones y condiciones experimentales controladas. Huelga decir que en los estudios de campo ese control resulta difícil y sin embargo hay que procurar aminorar los efectos de los diversos determinantes de la varianza de error. Ello puede lograrse en parte, dando instrucciones específicas y claras a los sujetos y excluyendo de la situación experimental los problemas ajenos a la investigación.

Es lo mismo acrecentar la confianza de las medidas que disminuir la varianza de error. A reserva de la exposición más amplia que se presenta más adelante, podemos decir que la confiabilidad es la "exactitud" de un grupo de puntuaciones. En la medida en que no fluctúan casualmente, es posible

(8) Véase a D. Campbell y J. Stanley, *Experimental and Quasi-Experimental Designs for Research*, Stokio: Ill.: Rand McNally, 1963, págs. 15, 49.

determinar su confiabilidad. Imaginemos un instrumento de medición nada confiable, uno que no nos permitiese en absoluto predecir la actuación futura de los individuos, que suministrase una ordenación jerárquica de una muestra de sujetos y otra ordenación totalmente distinta en otra aplicación de la misma prueba. Con semejante instrumento no sería posible identificar ni extraer varianzas sistemáticas, puesto que las puntuaciones obtenidas serían como números tomados de una tabla de números aleatorios. Este es el caso extremo. Imaginemos ahora diversos grados de confiabilidad y desconfiabilidad en las medidas de la variable dependiente. Cuanto más confiables sean las medidas, mejor identifiquemos y extraeremos las varianzas sistemáticas y menor será la varianza de error en relación con la varianza total.

Otra razón para reducir la varianza de error al máximo es permitir que las varianzas sistemáticas manifiesten su importancia, *si es que la tienen*. No podemos hacerlo si dicha varianza, y por ende el término de error, es demasiado grande. Cuando existe una relación procuramos descubrirla. Una manera de detectarla o de dejar que aparezca es encontrar las diferencias significativas que hay entre medias. Pero si la varianza de error es relativamente grande a causa de los errores incontrolados de la medición (y por lo mismo poco confiable), las varianzas sistemáticas —a las que antes dimos el nombre de “varianzas entre grupos”— no serán significativas. Así pues, aunque la relación exista, muy probablemente no será o no podrá ser descubierta.

El problema de la varianza de error se puede expresar por una sencilla fórmula matemática: recuérdese la ecuación:

$$V_t = V_b + V_e$$

donde  $V_t$  es la varianza total en un conjunto de medidas;  $V_b$  es la varianza entre grupos, que posiblemente se debe al influjo de las variables experimentales; y  $V_e$  es la varianza de error (en el análisis de varianza, la varianza entre grupos residual). Es claro que cuanto mayor sea  $V_e$ , menor deberá ser  $V_b$ , con determinada magnitud de  $V_t$ .

Más aún, examínense las siguientes ecuaciones:

$$t = \frac{\text{estadística}}{\text{error estándar de la estadística}}$$

y

$$F = \frac{V_b}{V_e}$$

Ambas expresan lo mismo: los denominadores han de ser medidas exactas del error aleatorio, a fin de que los numeradores de las fracciones de la derecha sean evaluados con precisión en busca de desviaciones notables respecto a expectativas azarosas.

Un ejemplo bien conocido ilustrará lo anterior. Recuérdese que en las discusiones sobre el análisis factorial de varianza y sobre el análisis de varianza de grupos correlacionados se hizo mención de la varianza ocasionada por la presencia de diferencias individuales en las medidas experimentales. Se dijo que si bien una buena aleatorización logra igualar a los grupos experimentales, habrá una varianza en las puntuaciones ocasionada por las diferencias individuales; por ejemplo, en las derivadas de la inteligencia, aptitud, etc. Ahora bien, en algunas situaciones, tales diferencias pueden ser muy notables. De ser así, la varianza de error y por consiguiente los denominadores de las ecuaciones  $t$  y  $F$ , serán “demasiado grandes” en comparación con los numeradores; vale decir, las diferencias individuales habrán sido distribuidas aleatoriamente entre dos, tres o cuatro grupos experimentales. A pesar de eso son fuente de varianza y por lo mismo aumentarán de manera excesiva la varianza entre grupos, la residual o la varianza típica (error estándar al cuadrado), es decir, los denominadores de las ecuaciones antes vistas.

### Guías de estudio

Hemos dicho que el diseño de investigación tiene por objeto obtener respuestas a las preguntas de investigación y controlar la varianza. Explique en detalle lo que significa esta afirmación. ¿En qué forma un diseño controla la varianza? ¿Por qué debe el diseño factorial controlar más varianza que el diseño unidireccional? ¿Cómo controla la varianza un diseño que utiliza sujetos apareados o medidas repetidas de los mismos sujetos? ¿Qué relación existe entre las preguntas de investigación e hipótesis y el diseño? Al contestar las preguntas, elabore un problema de investigación para ilustrar sus aseveraciones o utilice un ejemplo tomado del texto.

## Diseños inadecuados y criterios de los diseños

Todas las creaciones humanas tienen una ordenación o estructura: arquitectura, poesía, música, pintura, matemáticas, investigación científica. El ser humano concede suma importancia al contenido de sus creaciones, a menudo sin darse cuenta de que si carecen de una sólida estructura serán débiles y estériles, por muy rico e importante que sea su contenido.

Lo mismo sucede con la investigación científica. El hombre de ciencia necesita una forma viable y plástica con la cual expresar sus metas. Sin contenido —privado de una buena teoría y de buenas hipótesis y problemas— el diseño estará vacío. Pero sin forma, sin un armazón bien concebido y creado para alcanzar la finalidad que se persigue, no es posible lograr algo de valor. Por tanto, no es una exageración afirmar que muchos de los fracasos de la investigación de la conducta obedecen a la carencia de una estructura disciplinada e imaginativa.

El presente capítulo se centra sobre todo en los diseños de investigación inadecuados. Tales errores son tan comunes, especialmente en la educación, que merecen un estudio aparte, para que el estudiante pueda ser capaz de reconocerlos y comprender *por qué* son inadecuados. Este enfoque tiene una ventaja: el estudio de las deficiencias obliga a preguntar por qué algo es defectuoso, lo cual a su vez centra nuestra atención en los criterios con que se juzgan las suficiencias e insuficiencias. Por tanto, el estudio de esos diseños nos lleva a examinar los criterios del diseño de investigación. Aprovecharemos la oportunidad para describir el sistema simbólico que se empleará en este libro y para trazar una importante distinción entre la investigación experimental y la llamada investigación ex post facto.

### Método experimental y método ex post facto

Una exposición sobre el diseño ha de ir precedida por una diferenciación importante: la existente entre ambos métodos. En efecto, semejante distinción es tan importante que se le dedicará un capítulo más adelante. La locución latina *ex post facto* significa “a partir de lo ya acontecido”. Denota algo que sucede o se realiza *después* de un fenómeno que

tiene efecto retroactivo sobre el hecho. En este libro se usa por contraposición a “experimental” y se le ha dado un sentido específico y, ojalá, inequívoco.

Por experimento se entiende una investigación científica en la cual quien la realiza (el investigador) manipula o controla una o más variables independientes y observa la variable o variables dependientes en busca de una variación concomitante con la manipulación de las variables independientes. De lo dicho se deduce que el *diseño experimental* es aquél en que se *manipula* al menos una variable independiente. Hurlock manipuló los incentivos a fin de producir distintos grados de retención. Fleming y Antonnen manipularon la información del cociente intelectual dada a los maestros en el estudio dedicado a los efectos de las expectativas del maestro: a algunos se les dijo que ciertos alumnos poseían cocientes intelectuales muy superiores a los que realmente tenían.

En un *verdadero* experimento, el investigador tiene el poder de asignar los sujetos a los grupos experimentales. Lo ideal es que también posea el poder de *seleccionar* situaciones ideal a menudo se le niega. Si carece del *poder* de asignar los sujetos a los grupos experimentales o para asignar los tratamientos experimentales a los grupos, su estudio será un experimento mas no un *verdadero*.

En la investigación ex post facto es imposible manipular y asignar los sujetos o tratamientos por que la variable o variables independientes ya ocurrieron, por decirlo así. El investigador comienza con la observación de la variable dependiente y estudia retrospectivamente las variables independientes en busca de los efectos que pudieran tener en aquella. Cuando Getzels y Jackson compararon las características de niños “no creativos” de gran inteligencia y niños “creativos” de menor inteligencia, ya se habían formado los grupos “creativos y no creativos” y los grupos muy inteligentes y menos inteligentes.<sup>(1)</sup> Dicho con otras palabras, las variables independientes ya habían “ocurrido”. Getzels

(1) J. Getzels y P. Jackson, “Occupational Choice and Cognitive Functioning: Career Aspirations of Highly Intelligent and of Highly Creative Adolescents,” *Journal of Abnormal and Social Psychology*, LXI (1960), 119-125.