

CAPITULO 20

Diseño de investigación y sus aplicaciones: grupos aleatorizados

Es difícil decirle a una persona cómo llevar a cabo una investigación. A lo sumo, podemos tratar de cerciorarnos de que el principiante percibe las diversas posibilidades sería de su responsabilidad exclusiva. Al abordar un problema, el investigador deberá dejar que su mente divague, especule sobre las posibilidades e incluso que haga conjeturas. Una vez conocidas las alternativas, la etapa intuitiva del pensamiento enriquecerá la idea que tiene acerca de la investigación y lo conducirá más eficazmente a la etapa analítica de organización y estructuración del problema. Sólo entonces estará listo para planear su enfoque y decidir cuáles métodos analíticos y de investigación usará al poner en práctica sus ideas. Un buen diseño no es puro análisis, sino que también se necesita el pensamiento intuitivo porque ayuda a llegar a soluciones que salen de lo común. Quizá lo más importante: debe recordarse que el pensamiento intuitivo y el analítico dependen del conocimiento, la comprensión y la experiencia.

Este capítulo y el siguiente tienen como finalidad ante todo enriquecer las anteriores explicaciones del diseño y de orden estadístico con ejemplos concretos y sugerir las posibilidades fundamentales para estructurar la investigación, de manera que el estudiante logre resolver dichos problemas.

Diseño sencillo

con sujetos asignados aleatoriamente

En los capítulos 13 y 14 se comentaron e ilustraron los aspectos estadísticos del análisis de la varianza de una sola variable y factorial. El diseño implícito en la explicación será denominado aquí como *diseño con sujetos asignados aleatoriamente*. El paradigma del diseño general es el diseño 19-1:(1)

R	X	Y
	-X	Y

(1) De aquí en adelante, con la expresión "diseño general" se indicará el tipo de paradigma que aparece en esta página y explicado en los capítulos 18 y 19. La locución "análisis del diseño de varianza" demostrará el tipo de paradigma empleado en los capítulos precedentes, en éste y en el 21.

Ejemplos de investigación

Aunque parezca extraño, no abundan los ejemplos de este tipo de diseño. A continuación reseñamos tres de ellos. Dos se valieron de la asignación aleatoria, uno probablemente no. Es lamentable que algunos investigadores no mencionen cómo asignaron los sujetos a los grupos. A estas alturas debe ser evidente la necesidad de indicar el método con que se seleccionaron y asignaron los sujetos a los grupos experimentales.

Clark y Walberg: Recompensas masivas y lectura. Abundan las investigaciones con animales en que se estudian los efectos del refuerzo sobre el aprendizaje, y en cambio es escasa la experimentación controlada efectuada en el aula. Clark y Walberg, al trabajar con niños negros que cuatro años antes habían abandonado la escuela, se valieron de refuerzo masivo a fin de precisar su efecto en el aprovechamiento de lectura.(2) Un rasgo poco común fue la asignación aleatoria de los escolares a clases y de éstas al grupo experimental y al grupo de control. Al inicio del experimento, todos los alumnos fueron elogiados por su trabajo. Después que se estabilizaron los índices de recompensa por niño y por maestro (al cabo de seis sesiones), los maestros de los grupos experimentales recibieron instrucciones de duplicar o triplicar las recompensas, en tanto que a los de los grupos de control se les ordenó "que se mantuvieran al mismo nivel". Al terminar el período de tres semanas, todos los niños realizaron una prueba de lectura.

Las medias de ambos grupos fueron de 31.62 y 26.86, notablemente distintos al nivel de .01 ($F = 9.52$, a 1 y 108 grados de libertad).(3) A juzgar por el estudio, los niños pobres a quienes se refuerza en forma masiva obtienen puntuaciones muy superiores

(2) C. Clark y H. Walberg, "The Influence of Massive Reward on Reading Achievement in Potentially Urban Dropouts," *American Educational Research Journal*, V (1968), 305-310.

(3) Calcule el coeficiente intraclass de correlación, que fue de .13, que es la proporción de la varianza total causada por la variable independiente, o refuerzo masivo. Esto es substancial, en especial si se toma en cuenta la dificultad de la tarea y la escasez de recursos de Clark y Walberg. Advirtase que controlaron además la inteligencia; en el capítulo 21 examinaremos cómo lo hicieron. Un detalle importante e interesante de su estudio es que los administradores y maestros consideraron revolucionaria la idea del refuerzo.

res en lectura que aquellos que reciben un refuerzo moderado. Este estudio constituye un buen ejemplo de recursos sencillos aplicados con competencia y eficacia.

Wickens: Estudio sobre la generalización de estímulos y respuestas. Un interesante, ingenioso y valioso trabajo con importantes consecuencias para la educación es este estudio que efectuó Wickens por medio del condicionamiento.(4) Vale la pena comentarlo por dos razones interconexas: 1) evidentemente no se empleó la aleatorización y 2) es probable que se juzgara innecesaria. El segundo punto requiere una explicación. Con cualquier tipo de respuesta humana, en especial las de orden fisiológico, que sea universal en el *homo sapiens* y que no muestre una amplia gama de diferencias individuales, a veces no encierra ningún peligro omitir la randomización. Si todas las personas se asemejan muchísimo en una característica que se estudia, es obvio que carece de importancia recurrir o no a la asignación aleatoria. En este aspecto, cualquier individuo es representativo de la raza humana. Ejemplos de ello son la posesión de sangre, pulmones y latido cardíaco. Desde luego, el tipo de sangre y la frecuencia cardíaca y la acción de los pulmones cambian radicalmente el panorama cuando sirven de variables. En todo caso, en el estudio citado probablemente se supuso que todos los sujetos eran condicionables. Aun así, habría sido mejor asignar aleatoriamente los sujetos a los grupos, porque sabemos que se dan diferencias individuales en el condicionamiento o condicionabilidad. Y es posible que afectasen a los resultados del experimento.

Volvamos al experimento de Wickens. Las respuestas de los sujetos frente a una descarga eléctrica fueron condicionadas a un tono. Una vez terminado el condicionamiento, voltearon la palma de sus manos hacia arriba y se probaron sus respuestas condicionadas ante el primer tono y ante los demás. Se usaron tres grupos experimentales y uno de control. Los grupos I, II y III, o sea los experimentales, fueron condicionados y luego se les practicaron pruebas diferentes. Aunque no nos interesan los detalles, en general los sujetos fueron condicionados a un tono y sometidos a prueba con él y con los tonos una o dos octavas por encima o por debajo del tono original de condicionamiento. El número medio de respuestas (flexión) en una nueva posición de las manos reveló que hubo generalización de estímulos y respuestas. Las medias de los cuatro grupos fueron: I, 5.87; II, 3.69; III, 4.19; y en el grupo testigo, 0.16.

Con datos de los grupos I, II y III Wickens hizo un análisis de varianza simple, a fin de demostrar

(4) D. Wickens, "Studies of Response Generalization in Conditioning. I. Stimulus Generalization during Response Generalization," *Journal of Experimental Psychology*, XXXIII (1943), 221-227.

que no diferían en la respuesta promedio. La razón F , fue insignificante con lo que se demostró la generalización de estímulos, pues los sujetos respondieron en forma parecida al tono original y a los que diferían por una o dos octavas del original. Después comparó cada uno de los grupos experimentales con el grupo de control. Las comparaciones fueron significativas, lo que demostró que hubo generalización de respuestas, pues el grupo de control recibió descargas pero no fue condicionado y las respuestas fueron emitidas con la palma de la mano hacia arriba.

Golightly y Byrne: Formulación de actitudes como refuerzos positivos o negativos. ¿Se puede ayudar a que la gente aprenda discriminaciones si se refuerzan sus respuestas correctas con formulaciones de actitudes que concuerden con las de ellos? El refuerzo tradicional es: "sí-no", "correcto-incorreto", y expresiones análogas. Supongamos que cuando el sujeto emite una respuesta acertada en lugar de decir "correcto" e "incorrecto" decimos: "el partido demócrata es bueno" -y el sujeto es demócrata. ¿Actuarán como reforzadores estos enunciados de actitudes del mismo modo que lo hacen los granos con los pichones y la expresión "correcto-incorreto" con los niños? Golightly y Byrne comprobaron la idea valiéndose de tres grupos de 20 sujetos cada uno (asignados aleatoriamente), en el diseño 19-1 y de un análisis simple de varianza como marco de referencia.(5) El primero de los grupos, al que se sometió a la situación tradicional de premio-castigo, fue reforzado con las palabras "correcto" e "incorrecto" al emitir respuestas acertadas y erróneas. (La discriminación se hizo entre grande y pequeño; la primera palabra era correcta para la mitad de los sujetos y la segunda para la otra mitad.) Al segundo grupo se le reforzó con formulaciones de actitudes y se le puso en una situación donde se expresaban actitudes semejantes o contrarias a las del sujeto: si una discriminación era correcta, se presentaba una formulación con la cual coincidía; si era incorrecta, se presentaba una formulación con la cual discrepaba. El tercero era un grupo de control: sus miembros recibían formulaciones que manifestaban actitudes neutrales. Evidentemente se utilizaron comparaciones planeadas, y ambos grupos experimentales fueron comparados por separado con el grupo de control.

Ante todo, según lo previsto, el primer grupo fue superior a los dos restantes y lo que es más importante, el segundo fue superior al grupo de control: $F = 26.78$, $g_l = 1.54$, significativa al nivel de .001. Es evidente que los enunciados de actitudes son eficaces reforzadores en el aprendizaje discriminativo. Probablemente ello no causará gran sorpresa a muchas personas. Después de todo, nos dirán: si deseas

(5) C. Golightly y D. Byrne, "Attitude Statements as Positive and Negative Reinforcements," *Science*, CXLVI (1964), 798-799.

quedar bien con alguien concuerda con sus opiniones. No obstante, el asunto es mucho más complejo. En efecto, este estudio es teóricamente importante por combinar la teoría social y la del aprendizaje y por haber demostrado, en un experimento sumamente original, el poder reforzador de la formulación de opiniones. Además, es un buen ejemplo del diseño 19-1.

Diseños factoriales

El diseño general 19-1 continúa siendo fundamental, si bien la variación de la pauta de grupo experimental-grupo de control se altera aquí drásticamente por la adición de otros factores experimentales o variables independientes. Según una antigua definición del análisis factorial de varianza, *el diseño factorial es la estructura de investigación, en la que dos o más variables independientes se juxtaponen, a fin de estudiar sus efectos autónomos e interactivos en la variable dependiente.*

En un principio al lector le será un poco difícil situar el sistema factorial dentro del paradigma general del diseño 19-1, constituido por grupo experimental y grupo de control. No obstante, la generalización de la noción de grupo de control expuesta en el capítulo 19 debe haber aclarado las relaciones entre el diseño 19-1 y los diseños factoriales. Ahora reanudaremos la exposición. Tenemos las variables independientes A y B y la variable dependiente Y . El diseño factorial más sencillo, el 2×2 , ofrece tres posibilidades: A y B son variables activas; A es activa, B atributiva (o a la inversa); y A y B son variables atributivas. La última posibilidad es el caso ex post facto. Regresemos a la noción de grupo experimental-grupo de control: A puede dividirse en A_1 y A_2 (grupo experimental y grupo de control respectivamente), con la variable independiente adicional B fraccionada en B y B . Puesto que ya conocemos esta estructura, nos limitaremos a comentar uno o dos detalles metodológicos.

El procedimiento ideal para asignar a los sujetos es distribuirlos aleatoriamente en las cuatro casillas. Si A y B son variables activas, será posible y fácil; bastará dar arbitrariamente a los sujetos números de 1 a N (N = número total de sujetos) y luego mediante una tabla de números aleatorios escribir los números de 1 a N tal como aparecen en ella. Colóquelos en cuatro grupos conforme vayan apareciendo y después asigne los cuatro grupos de sujetos a las cuatro casillas. Para mayor seguridad, asigne los grupos a los tratamientos experimentales (las cuatro casillas) en forma aleatoria. Numere los grupos 1, 2, 3 y 4. A continuación extraiga esos números de una tabla de números aleatorios. Suponga que aparecen en el siguiente orden: 3, 4, 1 y 2. Coloque a los sujetos del grupo 3 en la casilla superior de la izquierda, los del grupo 4 en la casilla superior de la derecha, y así sucesivamente.

A menudo B será una variable atributiva: sexo, inteligencia, aprovechamiento, ansiedad, auto percepción, raza. Y entonces es preciso alterar la asignación de los sujetos. En primer lugar, como B es una variable atributiva no será posible asignar aleatoriamente a los sujetos a B_1 y B_2 . Si B es la variable "sexo", lo mejor que podemos hacer es distribuir aleatoriamente a los varones en las casillas $A_1 B_1$ y $A_2 B_1$, y luego proceder de igual manera con las mujeres. Con el caso ex post facto en que ambas variables son asignadas, naturalmente no se puede hacer la asignación aleatoria. Sin embargo, en rigor dudamos que se deba utilizar el análisis factorial de varianza con los datos de tipo ex post facto. Disponemos de una alternativa mucho más conveniente: el análisis de regresión múltiple, que estudiaremos en los capítulos 35 y 36.

Diseños factoriales con más de dos variables

A menudo se puede mejorar el diseño y depurar la información de una investigación por medio de la adición de grupos. En lugar de A_1 y A_2 , y B_1 y B_2 , se puede sacar provecho en un experimento de A_1 , A_2 , y A_3 , y B_1 , B_2 y B_3 . Los problemas prácticos y estadísticos se incrementan y en ocasiones llegan a ser muy difíciles conforme se agregan variables. Supongamos un diseño de tipo $3 \times 2 \times 2$ que cuenta con $3 \times 2 \times 2 = 12$ casillas, en cada una de las cuales se encuentran como mínimo dos sujetos y generalmente mucho más. (Es posible, aunque no muy sensato, tener un sujeto por casilla cuando se puede disponer de varios. Desde luego existen diseños que sólo tienen un sujeto por casilla.) Si decidimos que se necesitan 10 sujetos por casilla, deberán obtenerse $12 \times 10 = 120$ sujetos, mismos que se asignarán aleatoriamente. Con una variable más, el problema se vuelve más agudo y la manipulación práctica de la situación se torna también más difícil. Pero el manejo eficaz de este experimento nos permite verificar cierto número de hipótesis y aporta copiosa información. Las combinaciones de los diseños con tres, cuatro y cinco variables proporcionan una extensa variedad de diseños: $2 \times 5 \times 3$, $4 \times 4 \times 2$, $3 \times 2 \times 4 \times 2$, $4 \times 3 \times 2 \times 2$, etc.

Ejemplos de investigación con diseños factoriales

En el capítulo 14 se describieron ejemplos de diseños factoriales bidimensionales y tridimensionales. (Se recomienda que el lector los estudie de nuevo, ya que el razonamiento sobre el que descansa el diseño básico podrá entenderse ahora más fácilmente.) En el capítulo 14 se dieron varios ejemplos de diseños factoriales y, por tanto, los que vamos a reseñar se limitan a tres estudios que presentan rasgos poco comunes.

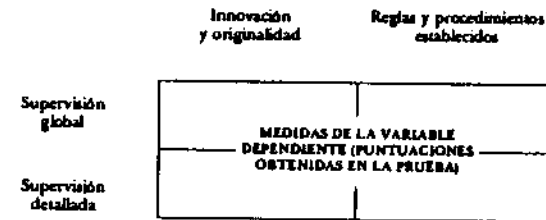


Figura 20-1.

Frederiksen, Jensen y Beaton: Ambiente organizacional y efectividad administrativa. El primer ejemplo es un diseño de tipo factorial 2×2 , concebido para probar ideas de gran interés acerca de las relaciones que hay entre los ambientes organizacionales y efectividad administrativa. (6) En la parte del estudio que bosquejamos —el estudio completo es un trabajo muy amplio— la variable independiente consistía en innovación y originalidad versus oposición a reglas y procedimientos establecidos, y la segunda era supervisión global versus supervisión minuciosa.

Ambas variables fueron manipuladas. Hubo algunas variables dependientes derivadas de un instrumento poco común (In Basket Test), entre ellas, cualidad general de la gestión administrativa, productividad, planes y discusiones. El diseño se muestra en la figura 20-1.

A pesar de que los resultados fueron desalentadores —poquísimas pruebas estadísticas fueron significativas— el diseño tiene importancia teórica y práctica. Adviértase la naturaleza singular de las variables independientes, las cuales fueron manipuladas mediante diversas instrucciones escritas que se dieron a los administradores que realizaron la prueba. Son escasos los estudios cuyas variables reflejan distintas ideas y prácticas administrativas y organizativas. También las variables dependientes revisten interés teórico y práctico, pues reflejan múltiples facetas de la conducta y del estilo directivo.

Lana: Efectos de la interacción de la preprueba. (7) Este caso es un estudio metodológico donde se utiliza el diseño Solomon comentado en el capítulo 19 en el diseño 19-6. Recuérdese que se trata de un diseño de cuatro grupos, con un grupo experimental y tres grupos de control, en el cual se aplican prepruebas al grupo experimental y al grupo de control. Lana sometió a prueba empírica las nociones de Solomon y Campbell. ¿Ejercen las prepruebas

un efecto sensibilizador sobre las actitudes? Recuérdese que una de las principales finalidades del diseño de Solomon era controlar posibles efectos de la preprueba en Y , la variable dependiente. Recuérdese asimismo que Solomon propuso, entre otros análisis estadísticos, un análisis factorial de tipo 2×2 para los cuatro conjuntos de las medidas de preprueba, a fin de separar la varianza de preprueba-posprueba y la varianza de las demás variables independientes. Lana adoptó estas ideas y las aprovechó.

Se aplicó al grupo I (el grupo experimental), y al grupo II (el primer grupo de control) una preprueba de actitudes. Los grupos III y IV no fueron sometidos a dicha preprueba. Una variable experimental era preprueba-no-preprueba. Otra era la comunicación-no-comunicación. A diferencia de la mayoría de los estudios actitudinales, esta última variable era menos importante que la variable "preprueba".

Tanto el diseño como una parte de los datos de Lana se muestran en el cuadro 20-1. (8) La diferencia entre las medias de comunicación y no-comunicación fue significativa al nivel .05. La diferencia promedio de preprueba-no-preprueba no fue importante, pero lo más notable es que la interacción no fue significativa. Por lo visto la preprueba no producía ningún efecto sensibilizador.

Si queremos entender este ingenioso estudio y sus hallazgos, hemos de saber perfectamente lo que Lana trataba de hacer. Intentaba probar la hipótesis de que la preprueba y el tratamiento interactúan. De acuerdo con la advertencia de Campbell referente al peligro de efectuar prepruebas, cabría esperar que los grupos que las hicieron, precisamente por eso, estarían sensibilizados para recibir la comunicación, no así los grupos a los que no se aplicó

(7) R. Lana, "Protest-Treatment Interaction Effects in Longitudinal Studies," *Psychological Bulletin*, LVI (1959), 293-300.

(8) N. Frederiksen, O. Jensen, y A. Beaton. *Organizational Climate and Administrative Performance*. Princeton, N. J.: Educational Testing Service, 1968.

(8) En realidad Lana se valió de cinco grupos. Su diseño y su análisis fueron más complejos de lo que se ha expuesto. Sin embargo, el diseño básico es el que se muestra. Las pruebas adicionales fueron de control.

la preprueba. Tal sensibilización, si existía, quedaría demostrada por una diferencia entre las medias de los grupos sometidos a la preprueba comunicación-no-comunicación, y no se manifestaría entre las medias de los grupos a los que no se les había aplicado esta preprueba. Es decir, en el cuadro 20-1, las medias de las dos casillas situadas en la parte superior diferirían notablemente, no así las medias de las casillas de la parte inferior.

Hoyt: *Conocimiento de los maestros y aprovechamiento de los alumnos.* (9) Este estudio fue planeado para contestar a una importante pregunta pedagógica de índole teórica y práctica: ¿qué efectos tiene en el aprovechamiento y en las actitudes de los escolares el que se dé a conocer a los profesores las características de sus alumnos? En un estudio previo Ojemann y Wilkinson se habían propuesto demostrar que la retroalimentación de conocimiento acrecienta el aprovechamiento y la personalidad del alumno. (10) En su estudio, Hoyt exploró varios aspectos afines de la cuestión esencial y se valió de un diseño reciente y de los avances estadísticos para aumentar la validez interna de la investigación. Recurrió a tres diseños factoriales: $2 \times 2 \times 3$; $2 \times 3 \times 3$; 3×3 ; el primero lo usó tres veces en tres diferentes asignaturas, el segundo y el tercero los empleó dos veces, una vez en cada uno de los sistemas escolares.

Se procederá a bosquejar la estructura del primer diseño. Las variables independientes estaban constituidas por tratamientos, nivel de capacidad, sexo y escuelas. Los tres tratamientos eran: no-información (N), calificaciones (T) y calificaciones junto con otra información (TO). Ninguno requiere explicación alguna. Fueron asignados aleatoriamente a los grupos experimentales: los niveles de capacidad fueron cociente intelectual alto, mediano y bajo. Las dos variables restantes, sexo y escuela, son obvias. Los alumnos de octavo grado fueron asignados aleatoriamente dentro de las poblaciones de sexo y aptitudes escolares. (11) El diseño se muestra en la figura 20-2. Como el diseño es complejo, conviene examinar el aspecto que presentará el análisis final de la tabla de varianzas y conocer las hipótesis que pueden ser sometidas a prueba. Antes, hay que señalar que los resultados referentes al aprovechamiento fueron en gran parte indeterminados (o negativos). Con excepción de una interacción, las

(9) K. Hoyt, "A Study of the Effects of Teacher Knowledge of Pupil Characteristics on Pupil Achievement and Attitudes towards Classwork," *Journal of Educational Psychology*, XLVI (1955), 302-310.

(10) R. Ojemann y F. Wilkinson, "The Effect on Pupil Growth of an Increase in Teachers' Understanding of Pupil Behavior," *Journal of Experimental Education*, VIII (1959), 143-147.

(11) Aunque no suele hacerse, aquí era apropiado. El procedimiento consiste en seleccionar y asignar dentro de cada sexo por separado y dentro de cada nivel de capacidad también por separado.

Cuadro 20-1. Diseño y datos del estudio de Lana sobre el efecto de la interacción de preprueba y tratamiento

	Preprueba	
	Preprueba	Sin preprueba
Comunicación	42.96	42.90
No comunicación	40.28	40.77

razones *F* fueron poco significativas. Un notable resultado positivo fue que las actitudes de los alumnos hacia los maestros parecían mejorar cuando aumentaban los conocimientos de estos últimos, hallazgo interesante e importante.

En el cuadro 20-2 se consignan las fuentes de varianzas y los grados de libertad empleados en la tabla de análisis de varianzas. ¡Un solo experimento produjo 14 pruebas estadísticas! Aunque, por supuesto, algunas de ellas carecen de importancia y podrían omitirse. Las de mayor importancia (marcadas con asteriscos en el cuadro) son las que incluyen la variable "tratamiento". La prueba más importante es la que se efectúa entre un tratamiento y otro: el primero de los efectos principales. El segundo lugar, aunque quizá tiene la misma importancia, corresponde a las interacciones que suponen tales tratamientos. Tomemos la interacción tratamientos \times sexo. Si fuera significativa, implicaría que la cantidad de información que posee un maestro sobre los alumnos influye en el aprovechamiento, pero que los varones no son influidos de igual manera que las mujeres. Los varones cuyos maestros poseen información sobre ellos, tendrán un mejor rendimiento que aquellos cuyos maestros no la tienen. Ocurre lo contrario con las mujeres o quizá no exista diferencia en ningún sentido. (12)

Las interacciones de segundo o triple orden son difíciles de interpretar. Probablemente es cierto que rara vez son significativas. Si lo son requieren de un estudio especial; una de las mejores formas de hacerlo es la representación gráfica (expuesta en un capítulo precedente). También son útiles las tablas de medias con divisiones cruzadas. El lector encontrará mayor orientación en el libro de Edwards. (13)

(12) Al lector le será útil tomarse la molestia de estudiar los cuadros de estas interacciones. Por ejemplo, la que estamos tratando sería:

	N	T	TO
mujer			
varón			

Anote las medias e interprete los resultados de varias combinaciones.

(13) A. Edwards, *Experimental Design in Psychological Research*, 2a. ed. Nueva York: Holt, Rinehart and Winston, 1960, caps. 11, 12.

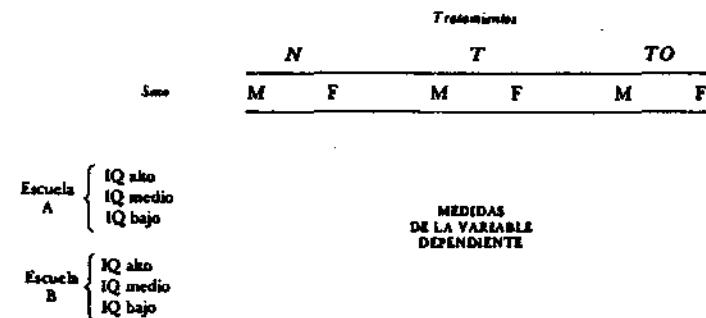


Figura 20-1.

Valoración de los diseños con sujetos asignados aleatoriamente

Todos estos diseños son variantes o extensiones del 19-1, o sea el diseño básico con grupo experimental y grupo de control, en los cuales los sujetos son asignados aleatoriamente a uno y otro grupos. Y por ello reúnen las ventajas del diseño básico, la más notable de las cuales es la de asignación aleatoria y la consecuente capacidad para suponer igualdad

preexperimental aproximada de los grupos experimentales en todas las variables independientes posibles. Se controlan historia y maduración porque transcurre poquísimos tiempo entre la manipulación de *X*, y la observación y medición de *Y*. No existe el riesgo de una contaminación debido a la preprueba. Otras dos ventajas de estos diseños, derivadas de las múltiples variaciones posibles, son la flexibilidad y la aplicabilidad. Se les puede usar para resolver

Cuadro 20-2. Fuentes de varianzas y grados de libertad en un diseño factorial de $3 \times 3 \times 2 \times 2$ con las siguientes variables: tratamientos, habilidad, sexo y escuela (se omiten el total de los grados internos de libertad)

Fuente	df
Efectos principales:	
*Entre tratamientos	2
Entre niveles de capacidad	2
Entre sexos	1
Entre escuelas	1
Interacciones de primer orden:	
*Interacción: tratamientos \times habilidad	4
*Interacción: tratamientos \times sexo	2
*Interacción: tratamientos \times escuela	2
Interacción: habilidad \times sexo	2
Interacción: habilidad \times escuela	2
Interacción: sexo \times escuela	1
Interacciones de segundo orden:	
*Interacción: tratamientos \times habilidad \times sexo	4
*Interacción: tratamiento \times habilidad \times escuela	4
Interacción: habilidad \times sexo \times escuela	2
Interacción de tercer orden:	
Interacción: tratamiento \times habilidad \times sexo \times escuela	4
Grados internos de libertad	
Total	

muchos problemas de investigación de la conducta, pues parecen adaptarse perfectamente a los tipos de diseño que requieren los problemas e hipótesis de carácter social, científico y pedagógico. Por ejemplo, los diseños de una sola variable pueden contener cualquier número de métodos, y la comprobación de estos últimos es una necesidad esencial en pedagogía. Las variables que siempre requieren control en la investigación de la conducta (sexo, inteligencia, aptitud, clase social, escuelas y tantas otras) pueden incorporarse en los diseños factoriales para así poder controlarlas. Con los diseños factoriales también es fácil tener mezclas de variables activas y atributivas; otra tarea de primer orden.

Por supuesto, también hay limitaciones. Una crítica que se hace a estos diseños es que no permiten las pruebas de igualdad de grupos, como lo hacen los diseños basados en las prepruebas y pospruebas. En realidad, no es una crítica válida por dos motivos: con un número suficiente de sujetos y con una buena asignación aleatoria, cabe suponer que los grupos sean iguales según hemos visto; y es posible comprobar la igualdad de los grupos en variables diferentes de Y , la variable dependiente. Los datos concernientes a variables como inteligencia, aptitud y aprovechamiento, generalmente se encuentran en los archivos escolares. Por ejemplo, es fácil efectuar un análisis de varianza de la variable inteligencia.

Otra dificultad es de carácter estadístico. Se debe tener números iguales de casos en las casillas de los diseños factoriales. (Es posible trabajar con cenes desiguales, pero es a la vez una torpeza y una amenaza contra la interpretación. Para evitar las pequeñas discrepancias se eliminan aleatoriamente algunos casos.) Ello impone una limitación al empleo de estos diseños, porque a menudo no se pueden tener números iguales en cada casilla. Los diseños de una sola variable con sujetos asignados aleatoriamente no son tan delicados; los números desiguales no constituyen un problema difícil. (14)

En comparación con los diseños de grupos apareados, los de grupos aleatorios suelen ser menos precisos; es decir, el término de error es ordinariamente mayor en igualdad de circunstancias. Sin embargo, no creemos que se trate de una cuestión de gran interés. En algunos casos sí lo es; por ejemplo, cuando se necesita efectuar una prueba muy exacta de una hipótesis. En la investigación de la conducta tal vez sea mejor considerar insignificante cualquier

efecto que no sea lo suficientemente poderoso como para hacerse sentir al mismo nivel o sobre el ruido aleatorio de un diseño de este tipo.

En conjunto, son diseños poderosos flexibles, útiles y de amplia aplicación. A nuestro juicio son los mejores diseños de uso general, quizá los primeros que han de tomarse en cuenta al planear el diseño de un estudio de investigación.

Guías de estudio

- Al estudiar el diseño de investigación, es útil hacer algunos análisis de varianza—tantos como sea posible: de una sola variable y factoriales con dos variables. Intente incluso un análisis con tres variables. Mediante este ejercicio estadístico el lector podrá comprender mejor los diseños. También podrá asignar nombres de variables a sus "datos", en vez de trabajar con números exclusivamente. A continuación se ofrecen algunas sugerencias provechosas para elaborar proyectos con números aleatorios.
 - Extraiga tres grupos de números aleatorios, del 0 al 9. Asigne nombres a las variables independientes y dependiente. Formule una hipótesis, traduzcala al lenguaje estadístico y de diseño. Efectúe un análisis unidireccional de varianza. Interprete.
 - Repita 1(a) con cinco grupos de números.
 - Agregue 2 a los números de cualesquiera de los grupos y reste 2 a los de otro grupo. Repita el análisis estadístico.
 - Extraiga cuatro grupos de números aleatorios, 10 en cada uno. Incorpórellos al azar al diseño factorial de tipo 2×2 . Realice un análisis de varianza.
 - Altere los números de las dos casillas de la derecha sumando 3 a cada uno. Repita el análisis. Compárelos con los resultados de 1(d).
 - Altere los números de los datos de 1(d) del modo siguiente: sume 2 a cada uno de los números que se encuentran en las casillas de la parte superior izquierda y de la parte inferior derecha. Repita el análisis. Interprete.
- Consulte la sugerencia 2 del capítulo 13. Se tienen tres grupos. ¿A qué diseño general pertenece este ejemplo? ¿Por qué? ¿Hay un grupo de control?
- Consulte las sugerencias 2 y 3 del capítulo 14. Vuelva a hacer los ejemplos. ¿Son fáciles ahora? Ponga mucha atención al 3(c).
- Suponga que usted es el director de una escuela primaria. Algunos de los profesores de cuarto y quinto grados quieren prescindir de los libros de trabajo. Al inspector no le agrada la idea, pero está dispuesto a que usted demuestre la opinión de que esos libros no tienen gran importancia.

(14) El lector que desee consultar un estudio más avanzado hallará una exposición sobre las n desiguales, el problema de desproporcionalidad y la manera de ajustar las desigualdades, en G. Snedecor y W. Cochran, *Statistical Methods*, 6a. edición, Ames, Iowa: Iowa State University Press, 1967, cap. 16. Generalmente un procedimiento mejor para manejar el problema consiste en utilizar el análisis de regresión múltiple (véase el capítulo 36). Probablemente los estudios complejos como el de Hoyt, antes citado, puedan manejarse mejor por medio del análisis de regresión múltiple.

Cuadro 20-3. Datos hipotéticos (medias) de un experimento ficticio de tipo factorial

Sexo	Métodos			
	A_1	A_2	A_3	
Varones	45	45	36	42
Mujeres	35	39	40	38
	40	42	38	

- (Uno de los maestros llegó a decir que producen efectos negativos tanto en ellos como en los alumnos.) Elabore dos planes de investigación y diseños para comprobar la eficacia de los libros de trabajo: un diseño simple y un diseño factorial. Considere las variables aprovechamiento, inteligencia y sexo. También puede considerar la actitud de los maestros hacia los libros de trabajo como una variable independiente.
- He aquí un difícil problema de diseño de investigación. Imagine que es usted el presidente de una comisión creada para valorar los efectos de un plan de estudios que se va a establecer a nivel de licenciatura. Los partidarios del plan dicen que es superior a los programas ordinarios en todos los aspectos: aprovechamiento, actitud del alumnado, etc. Trace un plan y un diseño de investigación para verificar la eficacia del nuevo programa. ¿Cuáles son las limitaciones del diseño? ¿Cuáles serán las limitaciones de cualquier diseño en un problema de esta índole?
 - Suponga que se efectuó una investigación donde se utilizaron los métodos de enseñanza y el sexo como variables independientes y el aprovechamiento como variable dependiente, que arrojó los resultados anotados en el cuadro 20-3. Los números en el interior de las casillas son medias ficticias. Las razones F de los métodos y del sexo no son apreciables. La razón F de la interacción es significativa al nivel .01. Interprete los resultados estadísticamente y en forma substancial. Para hacer esto último nombre los tres métodos.
 - Lea los dos estudios que se citan a continuación: uno en el campo de la psicopedagogía y otro en el campo de la psicología social. Ambas investigaciones importantes e interesantes, bien diseñadas y ejecutadas.

Davis, J., y Klausmeier, H. "Cognitive Style and Concept Identification," *Journal of Educational Psychology*, LXI (1970), 423-430.

Jones E. "Conformity as a Tactic of Ingratiation," *Science*, CXLIX (1965), 144-150.

- En su opinión, ¿demostraron los autores de manera adecuada sus hipótesis?
 - Una de las variables independientes del estudio de Davis y Klausmeier es la dependencia de campo (estilo cognitivo), que como se sabe está correlacionada con la inteligencia. Los autores no incluyeron la inteligencia como variable del estudio. ¿Crea ello problemas en la interpretación de los resultados? En caso afirmativo, ¿por qué?
 - Ordene de nuevo las tablas 1 y 2 del informe de Jones en tablas de medias de tipo 2×2 . Interpretelas.
- Suponga que un investigador efectuó un estudio factorial de tipo $3 \times 2 \times 2$. Presente el análisis final de la tabla de varianza, designando los tres efectos con letras A , B y C . Apunte en ella los grados de libertad. En seguida nombre tres variables independientes y una variable dependiente que se ajusten al modelo. Haga que A sea significativa, B no significativa, y C significativa y la interacción de $A \times B$ significativa. Interprete los resultados.
Nota: hay siete posibles pruebas estadísticas. Consulte el cuadro 14-9 del capítulo 14, pero tenga cuidado pues hay diferencias.