

Diseño de investigación y sus aplicaciones: grupos correlacionados

Detrás de todos los diseños de grupos correlacionados se encuentra un principio básico: (1) hay una varianza sistemática en las medidas de la variable dependiente, debida a la correlación entre los grupos en alguna variable relacionada con la variable dependiente. La correlación y su varianza concomitante pueden ser introducidas en las medidas, y en el diseño, de tres maneras: 1) usando las mismas unidades (sujetos, por ejemplo) en todos los grupos experimentales; 2) apareando las unidades en una o más variables independientes relacionadas con la variable dependiente; 3) utilizando en el diseño más de un grupo de unidades, tales como clases o escuelas. Pese a sus diferencias, estas tres maneras de introducir la correlación en las medidas de la variable dependiente son fundamentalmente idénticas. Procedemos a examinar las consecuencias que tiene en el diseño este principio básico y a exponer las tres maneras de aplicarlo.

Paradigma general

Con excepción de los diseños factoriales correlacionados y los llamados diseños anidados, todos los demás paradigmas del análisis de varianza en los diseños de grupos correlacionados se pueden bosquejar con facilidad. El paradigma general se muestra en la figura 21-1. Con objeto de resaltar las fuentes de varianza se indicaron las medias de columnas e hileras. Asimismo se insertaron las medidas (las Y) individuales de la variable dependiente. (2) Se advierte que son dos las fuentes de varianza sistemática:

(1) La palabra "grupo" debe tomarse como un conjunto de puntuaciones. Así entonces no habrá confusión al clasificar como diseño multigrupal al experimento que consta de varios intentos.

(2) Resulta útil conocer el sistema de subcritos que acompañan a los símbolos empleados en matemáticas y estadística. Una tabla rectangular de números se denomina *matriz*. Las entradas son letras o números. Si se utilizan letras se suele identificar cualquier entrada de una matriz particular con dos (a veces más) subíndices. El primero indica el número de *hileras*; el segundo, el de *columnas*. Por ejemplo, Y_{12} designa la medida de Y en la tercera hilera y en la segunda columna. Se acostumbra generalizar el sistema añadiendo subíndices literales. En este libro i simboliza cualquier número de hilera y cualquier número de columna. Cualquier número de la matriz se representa por Y_{ij} ; cualquier número de la tercera hilera, por $Y_{i.}$; y cualquier número de la segunda columna, por $Y_{.j}$.

La que se debe a las columnas y tratamientos, la debida a las hileras -diferencias individuales o de unidades. El análisis de varianza ha de ser del tipo de dos variables.

El razonamiento ilustrado en la figura 21-1 no le será difícil a quien haya estudiado el tema de correlación-varianza expuesto en el capítulo 15, donde se presentaron la estadística correspondiente y algunos problemas de los grupos correlacionados. Es evidente que la intención del diseño consiste en maximizar la varianza entre tratamientos, identificar la varianza entre unidades y minimizar la varianza de error (residual). El principio *maximizar* es aplicable también en este caso. La única diferencia entre los diseños de grupos correlacionados y de sujetos asignados aleatoriamente radica en la varianza de hileras o unidades.

Unidades

Las unidades usadas no alteran los principios de varianza en lo más mínimo. Es necesario entender este punto con claridad. La concepción habitual del diseño de grupos correlacionados parece basarse con bastante firmeza en el apareamiento de sujetos. La palabra "unidades" se usó deliberadamente en la figura 21-1 a fin de superar esta idea y substituir una concepción generalizada. En principio no importa lo que sea substituido por "unidades": "Sujetos", "pares", "clases", "escuelas", "zonas escolares", "ciudades" y "condados" son cosas que "las unidades" pueden suplir adecuadamente. Lo importante es saber si las unidades, sin atender a lo que sean, difieren entre sí. Si es así, se introduce la *varianza entre unidades*. En este sentido, hablar de grupos o sujetos correlacionados es lo mismo que hablar de varianza entre grupos o sujetos. La noción de diferencias individuales se extiende a las *diferencias de unidades*.

El auténtico valor de los diseños de grupos correlacionados parece residir en que no sólo capacitan al investigador para aislar y estimar la varianza resultante de la correlación, sino que le sirven de guía al diseñar una investigación para aprovechar las diferencias que ordinariamente existen entre unidades.

| Unidades | Tratamientos | | | | | |
|----------|--------------|----------|----------|-----|----------|------------|
| | X_1 | X_2 | X_3 | ... | X_k | |
| 1 | Y_{11} | Y_{12} | Y_{13} | ... | Y_{1k} | $M_{1.}$ |
| 2 | Y_{21} | Y_{22} | Y_{23} | ... | Y_{2k} | $M_{2.}$ |
| 3 | Y_{31} | Y_{32} | Y_{33} | ... | Y_{3k} | $M_{3.}$ |
| ... | ... | ... | ... | ... | ... | ... |
| n | Y_{n1} | Y_{n2} | Y_{n3} | ... | Y_{nk} | $M_{n.}$ |
| | $M_{.1}$ | $M_{.2}$ | $M_{.3}$ | ... | $M_{.k}$ | $(M_{.j})$ |

Figura 21-1.

Si un estudio de investigación contiene distintas clases de la misma escuela, dichas clases serán la posible fuente de la varianza. En consecuencia, será conveniente utilizarlas como unidades en el diseño. Las conocidas diferencias entre las escuelas son importantísimas fuentes de varianza en la investigación pedagógica. Se les puede manejar factorialmente o como se indicó en los diseños descritos en el presente capítulo. En efecto, si nos fijamos atentamente en un diseño factorial con dos variables independientes, una de ellas las *escuelas*, y en un diseño de grupos correlacionados con las unidades *escuelas*, nos encontraremos fundamentalmente ante el mismo diseño. Estudie la figura 21-2. A la izquierda se muestra un diseño factorial y a la derecha un diseño de grupos correlacionados. Parecen iguales, y efectivamente lo son en lo que respecta al principio de varianza. (Las únicas diferencias serían los números de puntuaciones colocados dentro de las casillas y el tratamiento estadístico.)

Diseño de ensayos

repetidos con un solo grupo

En este diseño, un mismo grupo recibe tratamientos distintos en diferentes momentos o es medido en momentos diversos. En un experimento de aprendizaje, al mismo grupo de sujetos se le dan varias tareas de diferente complejidad, o bien la manipulación experimental consiste en impartir principios de

aprendizaje en órdenes diversos; por ejemplo, de lo simple a lo complejo, de lo complejo a lo simple, de todo a la parte, de la parte al todo. No pocos estudios longitudinales pertenecen a esta categoría. Se han logrado buenos resultados en los estudios que utilizan este diseño para determinar el crecimiento físico y mental de los niños, no obstante sus defectos intrínsecos. El éxito seguramente se debe a la resistencia de los procesos evolutivos a las influencias externas que no sean masivas.

Dijimos antes que el mejor apareamiento posible de los sujetos es el del sujeto consigo mismo. También se mencionaron las dificultades inherentes al problema de control. Una de ellas nos recuerda la sensibilización ocasionada por la preprueba, que puede originar una interacción entre ésta y la variable manipulada experimentalmente. Otra estriba simplemente en que los sujetos maduran y aprenden con el transcurso del tiempo. El sujeto que haya pasado por uno o dos ensayos en una manipulación experimental y se encuentra ante otro, no es igual al que haya pasado solamente por uno. Las situaciones experimentales difieren muchísimo, por supuesto. En algunas situaciones, acaso la repetición de ensayos no influya demasiado en el rendimiento de los sujetos durante intentos posteriores; en otras situaciones, tal vez sí. El problema de cómo los individuos aprenden o desarrollan una excesiva sensibilización durante el experimento es difícil de resolver.

| | Tratamientos | | Escuelas | Tratamientos | |
|------------------|--------------|-------|----------|-----------------------------------|-------|
| | A_1 | A_2 | | A_1 | A_2 |
| | B_1 | | 1 | | |
| Escuelas | B_2 | | 2 | | |
| | B_3 | | 3 | | |
| Diseño factorial | | | | Diseño con grupos correlacionados | |

Figura 21-2.

| Pares | Tratamientos | | |
|-------|--------------|----------|-------|
| | X_1 | X_2 | |
| 1 | Y_{1a} | Y_{1b} | M_1 |
| 2 | Y_{2a} | Y_{2b} | M_2 |
| 3 | Y_{3a} | Y_{3b} | M_3 |
| . | . | . | . |
| . | . | . | . |
| n | Y_{na} | Y_{nb} | M_n |
| | M_a | M_b | |

Figura 21-3.

En definitiva, la historia, la maduración y la sensibilización son posibles deficiencias de la repetición de ensayos. El efecto regresivo puede ser también una deficiencia porque, según se vio en un capítulo anterior, los que obtienen puntuaciones bajas tienden a conseguir puntuaciones altas y los que obtienen puntuaciones altas tienden a obtener puntuaciones inferiores en la repetición de la prueba, simplemente por la imperfecta correlación de los grupos. Es obvio que se necesita un grupo de control.

Pese a las dificultades fundamentales que supone el paso de tiempo, a veces este diseño es útil. Ciertamente es el diseño implícito en los análisis *ex post facto* de los datos "temporales", por ejemplo, si tenemos una serie de mediciones del crecimiento de niños, las diferentes épocas en que se efectuaron con esponderán a los tratamientos. El paradigma del diseño es idéntico al de la figura 21-1. Basta substituir a los "sujetos" por "unidades" y designar los "intentos" con X_1, X_2, \dots

De este paradigma general pueden derivarse casos especiales. El más sencillo es el diseño de un grupo con preprueba y posprueba, o sea el diseño 18-2(a), donde un grupo de sujetos recibió un tratamiento experimental precedido por una preprueba y seguido de una posprueba. Ya hemos mencionado las deficiencias de este diseño, por lo cual no necesitamos detenernos en él. Sin embargo, cabe señalar que especialmente en su forma *ex post facto* se asemeja muchísimo a la observación y al pensamiento debidos al sentido común. Una persona observa las prácticas pedagógicas de hoy y decide que no son muy buenas. A fin de formular su juicio, implícita o explícitamente, compara las prácticas actuales con las de antaño. Conforme a sus predisposiciones personales, entre varias causas posibles escogerá una o más que expliquen lo que considera el estado desastroso de la educación: "educación progresiva", "educadores", "degeneración moral", "carencia de principios religiosos sólidos", etc.

Diseños de dos grupos, con grupo experimental-grupo de control

Este diseño tiene dos formas, la mejor de las cuales X , que repetimos ahora, se describió en el diseño 19-2 del capítulo 19.

| M_r | X | Y (Grupo experimental) |
|-------|----------|------------------------|
| | $\sim X$ | Y (Grupo de control) |

En este diseño, primero los sujetos son apareados y después asignados aleatoriamente a ambos grupos. En la otra forma, se les aparean pero no son asignados de manera aleatoria. El segundo diseño puede indicarse con sólo suprimir el suscrito r de M_r (diseño descrito en el capítulo 18 en el diseño 18-4, como uno de los inadecuados).

El paradigma estadístico y de diseño, de este caballo de batalla, se muestra en la figura 23-1. La inserción de los símbolos de las medias designa las (dos) fuentes de varianza sistemática: *tratamientos y pares, columnas e hileras*; esto contrasta de manera evidente con los diseños de asignación aleatoria que se describieron en el capítulo 20 y cuya única varianza sistemática eran los tratamientos o columnas.

La variante más común de este diseño es la de dos grupos, con preprueba y posprueba. Véase el diseño 19-3(b). El paradigma estadístico y de diseño así como su fundamento se tratan más adelante (véase cuadro 21-2).

Ejemplos de investigación con diseños de dos grupos

Se han publicado centenares de estudios con dos grupos del tipo de sujetos apareados. En muchos de ellos, probablemente la mayoría, se usan sujetos apareados, que no han sido asignados aleatoriamente al grupo experimental ni al grupo de control. Con frecuencia es imposible saber si los investigado-

res usaron o no la aleatorización, puesto que en sus informes suelen omitir la forma en que distribuyeron los sujetos en los grupos. Los estudios descritos en los siguientes párrafos fueron escogidos, no sólo porque ilustran el diseño de grupos correlacionados, el apareamiento y los problemas de control, sino porque tienen importancia desde los puntos de vista histórico, educativo y psicológico. Dos de ellos son de gran alcance.

Estudio de Thorndike sobre la transferencia de instrucción

En 1924, Edward L. Thorndike publicó un excelente estudio sobre el supuesto efecto que tienen ciertas asignaturas en la inteligencia. (3) Los alumnos fueron apareados según sus puntuaciones obtenidas en la forma A mediante la cual se midió la variable dependiente: inteligencia. Este test sirvió además de preprueba. La variable independiente era "estudiar durante un año materias como historia, matemáticas y latín. Una posprueba, forma B de la prueba de inteligencia, que se aplicó al final del año. Thorndike recurrió a un ingenioso dispositivo para separar el efecto diferencial de cada asignatura al aparear en la forma A a los que estudiaban, por ejemplo: inglés, historia, geometría y latín con los que cursaban inglés, historia, geometría y taller pedagógico. Y así pudo comparar los efectos diferenciales del latín y del taller pedagógico en esos dos grupos. Los incrementos en las puntuaciones finales de la inteligencia fueron considerados un efecto conjunto del crecimiento y de la asignatura estudiada.

Pese a sus deficiencias fue un estudio colosal. Thorndike reconoció la falta de controles apropiados, como lo revela en el siguiente pasaje donde habla de los efectos de la selección:

La razón principal por la que los buenos pensadores parecen a primera vista ser fruto de ciertas asignaturas es que han realizado esos estudios... Cuando cursaron griego y latín, estas materias parecían producir buenos pensadores. Ahora que cursan física y trigonometría también estas materias parecen producir buenos pensadores. Si los alumnos más capaces estudiaban educación física y arte dramático, se debería en apariencia a estas asignaturas la formación de buenos pensadores. (4)

Thorndike señaló el modo de controlar la investigación pedagógica, lo cual ha hecho disminuir las explicaciones metafísicas y dogmáticas en educación. Su trabajo asestó un fuerte golpe a la teoría de la educación mental, que compara la mente con un rastrillo de afeitar que puede ser afilado suavizándolo con temas "difíciles".

(3) Thorndike, "Mental Discipline in High School Studies," *Journal of Educational Psychology*, XV (1924), 1-22, 83-98.

(4) *Ibid.*, pág. 98.

No es fácil valorar un estudio como el citado, cuya extensión e ingenio resultan impresionantes. No obstante, dudamos de la suficiencia de la variable dependiente: "inteligencia" o "capacidad intelectual". ¿Pueden las asignaturas cursadas durante un año ejercer un profundo efecto sobre la inteligencia? En la actualidad se opina que no es necesario un número tan grande de sujetos, y lo que es más importante, el de Thorndike no fue un "verdadero" experimento. En rigor, se trata de un estudio *ex post facto* en el que naturalmente no era posible la asignación aleatoria. Thorndike midió la inteligencia de los alumnos y dejó que operasen las variables independientes: las asignaturas. Como acabamos de mencionar, sabía que su experimento adolece de esta deficiencia de control. Aun así, es un clásico que merece respeto y un examen cuidadoso pese a sus deficiencias en la historia y en la selección (la maduración sí fue controlada).

El estudio de Ocho Años

Otro ambicioso intento de dar respuesta empírica a una importante pregunta pedagógica es el Estudio de Ocho Años, (5) cuya finalidad fue responder a la pregunta: ¿qué cualidades tienen los métodos "modernos" de educación en la enseñanza media en comparación con los métodos "tradicionales" con los que se prepara a los jóvenes para la universidad? En él se incluyeron 30 escuelas de enseñanza media superior situadas en distintas partes del país. Los alumnos que las frecuentaban estudiaron conforme a un plan fundamental de estudios. Luego que ingresaron a la universidad fueron igualados con los procedentes de escuelas "tradicionales" en sexo, inteligencia y otras variables. Se comparó el rendimiento de los dos grupos. En términos generales, los resultados revelaron que los egresados de escuelas progresistas lograban un rendimiento ligeramente superior a los que provenían de escuelas tradicionales.

Este estudio a gran escala, igual que el de Thorndike, es una investigación *ex post facto* con graves limitaciones de control, probablemente la más importante de las cuales sea la autoselección. Pese al apareamiento, acaso intervinieron otras variables cruciales que en el momento de iniciar el estudio, distinguieron a los alumnos procedentes de escuelas progresistas de los que provenían de escuelas tradicionales. Por ejemplo, ¿hay una diferencia notable entre los niños que asisten a las escuelas progresistas de enseñanza media superior y los que asisten a escuelas tradicionales? ¿Propician los padres de los primeros mejores actitudes frente al aprendizaje que los padres de los segundos? (Algunas de ellas eran planteles de prácticas para los universitarios.) Este estudio está sujeto a críticas y no obstante sigue siendo un destacado intento por obtener datos obje-

(5) W. Aikin, *The Story of the Eight-Year Study*. New York: Harper and Row, 1942.

tivos sobre complejas y difíciles cuestiones de carácter pedagógico.(6)

Estudio de Bandura y Menlove sobre la modelación y extinción

Damos un salto de varios años hasta llegar al reciente uso de medidas de preprueba y posprueba. Cada día es más difícil hallar estudios psicológicos que recurran al apareamiento. Pero quizá por la naturaleza de tantos problemas que entrañan cambio cada día aumenta el número de investigadores que utilizan pruebas y diseños relativamente complejos, con medidas de preprueba y posprueba. Escogemos uno de esos estudios, no sólo porque se vale de prepruebas y pospruebas, sino porque pertenece a una serie de magníficas investigaciones dedicadas a los problemas del aprendizaje y de la modelación sociales, que siguen orientaciones teóricas y prácticas similares.(7)

Bandura y Menlove midieron primero en 48 preescolares la conducta de evitación de animales (la preprueba). Los pequeños fueron asignados aleatoriamente a tres grupos experimentales: un solo modelo, modelo múltiple y grupo de control. En la primera situación los sujetos vieron una película en que un niño de cinco años paulatinamente se acercaba a un perro de aguas. En la segunda situación observaron a niños y niñas que interactuaban positivamente con perros de distintas razas. En la tercera situación se les proyectaron otras películas cuyo tema era ajeno a la investigación. La conducta de evitación fue sometida a prueba al final del tratamiento (posprueba) y otra vez en un estudio de seguimiento.

Los análisis utilizados se salían de lo común: a causa de la anomalía de distribuciones de las puntuaciones, se aplicaron pruebas estadísticas no paramétricas. En primer lugar, a fin de comprobar el efecto de *tendencia*—esto es, las diferencias de las puntuaciones en la preprueba, posprueba y estudio de seguimiento— se empleó un análisis Friedman no paramétrico. χ^2 era de 15.79, significativa al nivel .001. Recuérdese que la prueba de Friedman toma en cuenta la correlación. En este estudio se comprobó las diferencias que existían entre los tratamientos experimentales aplicados a los mismos niños. Es claro que el cambio fue notable. Otras pruebas demostraron que cambiaban las dos situaciones mode-

ladoras, no así la situación de control. El valor de las diferencias de los tres tratamientos experimentales fue sometido a prueba por medio del análisis de varianza de una variable no paramétrica, ideado por Kruskal-Wallis. Se analizaron las puntuaciones de cambio o diferencia obtenidas en la preprueba y el estudio de seguimiento. H era de 5.01, significativa al nivel .05. Otros análisis demostraron la eficacia de los tratamientos de modelación.

Los autores introdujeron entonces un ingenioso giro en el estudio: al final del experimento propiamente dicho aplicaron el procedimiento de modelación múltiple a los sujetos del grupo de control. Estos sujetos ya habían tenido tres puntuaciones de aproximación: preprueba, posprueba, estudio de seguimiento. Llamemos "posterapia" al cuarto conjunto de puntuaciones. Con la prueba de Friedman se determinó la importancia de las diferencias de los cuatro conjuntos. Obsérvese cuidadosamente que cada sujeto recibía cuatro puntuaciones y en consecuencia nos hallamos ante grupos correlacionados. χ^2 era de 13.42, significativo al nivel .01. Y así el procedimiento modelador también fue eficaz en los sujetos del grupo de control.

Para que quede perfectamente claro lo que se hizo, los tres principales análisis de Bandura y Menlove se sintetizan en el cuadro 21-1, con un simbolismo ya conocido o al menos evidente. No se olvide que la prueba de Kruskal-Wallis es un análisis de una sola variable y contiene rangos de todas las puntuaciones, en tanto que la prueba de Friedman es un análisis de dos variables y usa por separado los rangos de cada hilera. Así pues, Bandura y Menlove se valieron de los diseños 19-1 y 19-3, y escogieron tests estadísticos para cumplir con las exigencias de sus diseños. Se recomienda al lector que consulte el estudio original por su refinamiento teórico y práctico, así como por su gran interés e importancia práctica para la psicología y la educación. Al hacerlo, tenga a la vista el cuadro 21-1. (El informe original ni es fácil de entender ni está completo.)

Diseños multigrupales con grupos correlacionados

Varianza de unidades

Aunque es difícil aparear tres o cuatro conjuntos de sujetos y aunque en la investigación de la conducta ordinariamente no es posible ni deseable utilizar los mismos sujetos en cada grupo, hay situaciones naturales en que existen grupos correlacionados y que tienen particular importancia en la investigación pedagógica. Hasta hace poco, las varianzas resultantes de las diferencias entre clases, escuelas, sistemas escolares y demás unidades "naturales" no habían sido reguladas debidamente ni empleadas con frecuencia en el análisis de datos. Quizá la

Cuadro 21-1. Esquema del análisis del estudio de Bandura y Menlove sobre la modelación

| I. Tendencias | | | | II. Situaciones experimentales* | | |
|--|------------|------------|------------------------|---------------------------------|----|---|
| Sujetos | Preprueba | Posprueba | Estudio de seguimiento | IM | MM | C |
| 1 | Y_{1pr} | Y_{1po} | Y_{1su} | | | |
| 2 | Y_{2pr} | Y_{2po} | Y_{2su} | | | |
| . | . | . | . | | | |
| N | Y_{Npr} | Y_{Npo} | Y_{Nsu} | | | |
| Prueba de Friedman | | | | Prueba de Kruskal-Wallis | | |
| Puntuaciones del cambio (estudio de seguimiento-preprueba) | | | | | | |
| III. Tendencias de control | | | | | | |
| Sujeto | Preprueba | Posprueba | Estudio de seguimiento | Poste-rapia | | |
| 1 | Y_{1pr} | Y_{1po} | Y_{1su} | Y_{1ra} | | |
| 2 | Y_{2pr} | Y_{2po} | Y_{2su} | Y_{2ra} | | |
| . | . | . | . | . | | |
| 12 | Y_{12pr} | Y_{12po} | Y_{12su} | Y_{12ra} | | |
| Prueba de Friedman | | | | | | |

* IM = modelación individual; MM = modelación múltiple; C = control.

primera indicación de la importancia que tiene esta clase de varianza apareció en el excelente libro de Lindquist sobre el análisis estadístico en la investigación pedagógica.(8) En él, Lindquist concede gran valor a la *varianza de escuelas*. Las escuelas, clases y demás unidades educacionales tienden a diferir mucho en rendimiento, inteligencia, aptitudes y otras variables. El investigador de la educación debe estar atento ante esas *diferencias de unidades*, lo mismo que ante las diferencias individuales.

Examinemos un ejemplo obvio. Supongamos que un investigador escoge una muestra de cinco escuelas por su variedad y heterogeneidad; huelga decir que busca validez externa: la representatividad. Efectúa la investigación con alumnos de cinco escuelas y combina las medidas de ellas para comprobar las diferencias promedio en alguna variable dependiente. De ese modo deja de lado la varianza causada por las diferencias de las escuelas. Es comprensible que las medias difieran poco; la varianza de escuelas se entremezcla con la de error.

Pueden surgir errores crasos si se pasa por alto la varianza de escuelas. Uno de ellos consiste en escoger escuelas y asignar unas al grupo de control y

otras al grupo experimental. En este caso, la varianza entre ellas se fusiona con la de la variable experimental. Asimismo las clases, zonas escolares y otras unidades educacionales difieren y, por tanto, generan varianza. Es preciso identificar y regular las varianzas mediante el control experimental o estadístico, o por medio de ambos.(9)

Ejemplo hipotético de varianza de escuelas

Al realizar un experimento en varias escuelas, el mejor procedimiento es procurar reproducir el experimento en *cada una*. El procedimiento es más o

(9) A fin de que el estudiante esté preparado para efectuar investigaciones en situaciones escolares, necesita hacer un examen más amplio del diseño y de los problemas estadísticos en cuestión de lo que permite la extensión de este libro. Por ejemplo, es preciso conocer las diferencias entre las situaciones en que los alumnos pueden ser asignados aleatoriamente a los tratamientos experimentales y aquellas en que se deben usar clases intactas. A tal efecto recomendamos las tres obras: E. Lindquist, *op. cit.*, caps. IV y V, especialmente págs. 104-132, 145-163; E. Lindquist, *Design and Analysis of Experiments in Psychology and Education*, Boston: Houghton Mifflin, 1953, caps. 7 y 8; A. Edwards, *Experimental Design in Psychological Research*, 3a. ed. Nueva York: Macmillan and Winston, 1968, cap. 9. El tipo de diseño que hemos visto recibe también el nombre de *diseño de bloques de asignación aleatoria o diseño de réplicas aleatorias*.

(6) El estudio abrió además nuevos horizontes en la modelación. Los investigadores idearon instrumentos al darse cuenta que los educadores progresistas hacían hincapié en cualidades no medidas por las pruebas ordinarias de inteligencia y aprovechamiento. Los más importantes fueron las pruebas para medir la capacidad para pensar. Véase a E. Smith y otros, *Appraising and Recording Student Progress*. Nueva York: Harper and Row, 1947.

(7) A. Bandura y F. Menlove, "Factors Determining Vicarious Extinction of Avoidance Behavior Through Symbolic Modeling," *Journal of Personality and Social Psychology*, VIII (1968), 99-108.

(8) E. Lindquist, *Statistical Analysis in Educational Research*. Boston: Houghton Mifflin, 1940.

| Niveles (dispositivos, tipos, etc.) | Unidades | Medidas (tratamientos) | | |
|--|----------|------------------------|----------------|----------------|
| | | A ₁ | A ₂ | A ₃ |
| B ₁ | 1 | | | |
| | 2 | | | |
| | 3 | | | |
| | 4 | | | |
| | 5 | | | |
| B ₂ | 1 | | | |
| | 2 | | | |
| | 3 | | | |
| | 4 | | | |
| | 5 | | | |

Figura 21-4.

menos el siguiente. Ante todo, seleccione aleatoriamente m escuelas de n escuelas en alguna población bien definida (las escuelas primarias de la zona o el condado). Realice el experimento en cada una. Absténgase de establecer una escuela como grupo experimental aquí y otra como grupo de control allá, porque entonces se confunde la varianza que existe entre ellas. Luego proceda a analizarla. Un paradigma estadístico y de diseño se asemejaría al de la figura 21-1, con la excepción de que podría haber medias en lugar de puntuaciones individuales en el interior de las casillas. (También podría haber en ellas medias de diversas clases o puntuaciones individuales.)

Las situaciones escolares varían y en consecuencia quizá no sea posible aplicar todos los tratamientos experimentales en cada plantel. En tales casos es posible, aunque mucho menos conveniente, muestras al azar a las escuelas y asignarlas aleatoriamente a los tratamientos experimentales. (10) Aunque no sea posible seleccionar a las escuelas de esa manera, se puede realizar un experimento útil.

Diseños factoriales con grupos correlacionados

No cabe duda que los modelos factoriales se combinarán con la noción de unidades para producir un

(10) Ver a Linnquist, *Design and Analysis of Experiments in Psychology and Education*, cap. 7 en las págs. 187-188. En esta obra nos ofrece un buen ejemplo de investigación.

diseño de gran utilidad: el *diseño factorial con grupos correlacionados*. Es claro que cuanto más complejos se tornen éstos, mayor dificultad habrá en los problemas estadísticos e interpretativos. No obstante, no debemos permitir que tales dificultades nos impidan usarlos. A fin de demostrar su utilidad, basta introducir en el diseño de la figura 21-1 los niveles de sexo, inteligencia, o cualquier otra variable pertinente. Desde el punto de vista conceptual, ello no entraña dificultad alguna (la figura 21-4 muestra el aspecto que tendría ese diseño).

Sus ventajas y desventajas son similares a las de los diseños factoriales más complejos. La ventaja principal es su capacidad de aislar y medir las varianzas y verificar las interacciones. Nótese que es posible valorar las dos grandes fuentes de varianza, *tratamientos* (a) y *niveles* (b), lo mismo que la *varianza de unidades* (o sea, las diferencias entre a y b); asimismo se puede comprobar la importancia de las medidas de las *unidades*. Se pueden verificar además tres interacciones: *tratamientos por niveles*, *tratamientos por unidades* y *niveles por unidades*. Si en lugar de las medias se anotan las puntuaciones individuales dentro de las casillas, también se podrá demostrar la interacción triple. Obsérvese lo importante que puede ser semejante interacción, tanto en teoría como en la práctica. Por ejemplo se contestan preguntas como las siguientes: ¿funcionan los tratamientos en forma distinta con unidades diferentes? ¿Funcionan en forma distinta ciertos métodos con niveles diferentes de inteligencia, o con uno y otro

sexo, o con niños que no tengan el mismo nivel socioeconómico? (11)

Análisis de la covarianza

La invención de este análisis se debe a Ronald Fisher, y tiene extraordinaria trascendencia potencial en la investigación pedagógica y psicológica. Muchas veces hay que estudiar a los grupos tal como son; es imposible aparear los sujetos o asignarlos aleatoriamente. El análisis de covarianza viene entonces en ayuda del investigador. Es un ejemplo del uso creativo de los principios de varianza, comunes a la teoría de correlación y al análisis de varianza, que nos permiten resolver un antiguo problema de carácter analítico. En esencia, Fisher extendió al análisis de la covarianza su noción básica de dividir la varianza total de un conjunto de medidas en varianzas sistemáticas y de error.

Análisis de covarianza es una forma del análisis de varianza que prueba la significancia de las diferencias entre las medias de los datos finales del experimento, para lo cual toma en cuenta la correlación entre la variable dependiente con una o más covariadas, y ajusta las diferencias iniciales de las medias en los grupos experimentales.

Veamos el caso de un equipo de sociólogos, psicólogos y educadores que desean valorar los efectos que un programa terapéutico masivo ejerce sobre la lectura de niños pobres. Sabedores de que la inteligencia es una variable impredecible en cualquier estudio de lectura, desean controlarla. Cuentan con un grupo experimental y un grupo de control de niños pobres; los niños son asignados aleatoriamente. (12) Disponen asimismo de las puntuaciones que cada uno obtuvo en una prueba de inteligencia y las emplean en un análisis de covarianza. Dicho análisis corregirá las diferencias iniciales de intelli-

(11) Al lector que haya cursado estudios avanzados de psicología y pedagogía seguramente le interesará saber cómo manejar la varianza de unidades (escuelas, clases, etc.) en diseños factoriales de mayor complejidad. Edwards ofrece instrucciones detalladas, *op. cit.* cap. 13.

(12) La dificultad práctica de la asignación aleatoria en una situación como esta, aunque considerable, quizá se ha exagerado. Por lo menos hay que procurar tener un grupo de control y un grupo experimental con niños asignados aleatoriamente. A estas alturas el lector comprenderá la dificultad que encierra valorar correctamente los efectos de un programa sin la comparación que proporcionan estos grupos y la asignación aleatoria. Por desgracia, demasiados investigadores suponen, muchas veces sin un buen fundamento, que no es posible establecer las condiciones apropiadas para una investigación. Siempre hay que intentar establecer aquellas que nos permitan extraer inferencias válidas de los datos, alterando las condiciones únicamente después de tratar de formularlas. Es probable que no pocas valoraciones pedagógicas están viciadas por suposiciones erróneas antes de iniciarse. En el caso presente, administradores, maestros y padres están seguros de que los niños del grupo de control recibirán tratamiento terapéutico al terminar el experimento. Lo que bien podría denominarse "deterioro por una suposición errónea" ha arruinado tantas valoraciones de programas educativos como la ignorancia.

gencia mostradas por ambos grupos, para lo cual tomará en cuenta la correlación de las medidas de inteligencia y lectura. En otras palabras, el procedimiento prácticamente controla la inteligencia en forma muy parecida a como lo hace el apareamiento. En realidad, suprime la varianza ocasionada por la inteligencia en las medidas de la variable dependiente antes de aplicar la prueba de significancia. Las sumas de los cuadrados y los cuadrados promedio "ajustados" a la covarianza se emplean en el análisis final de la tabla de varianza. Además, el análisis de covarianza fácilmente arroja tres estimaciones de la correlación de las medidas de inteligencia y lectura, una de las cuales es la mejor estimación de la "verdadera" correlación entre tales medidas.

Aun si los investigadores carecieran de medidas de inteligencia, podrían usar las mediciones de lectura de la preprueba en la misma forma. El análisis de covarianza deberá utilizarse para analizar las medidas finales de significancia, pero también deberá ajustarse de acuerdo a las diferencias de preprueba que existan entre los grupos. Si no se violan gravemente las suposiciones en que descansa dicho análisis, el método puede aplicarse en gran parte de la investigación social, científica y pedagógica, particularmente en la última, porque las medidas de las variables importantes que sirven de covariadas se encuentran disponibles en los archivos de escuelas, colegios, universidades, etc.

Descripción de los procedimientos seguidos en el análisis de covarianza

La covarianza (13) fue definida en el capítulo 6 como el promedio de los productos cruzados de las puntuaciones de desviación de dos variables, X y Y . (Esa puntuación está definida por $x = X - M_x$, o $y = Y - M_y$.) Si se tienen las puntuaciones de X y Y , de varios individuos obtenidas en una prueba de inteligencia y lectura, nos hallaremos ante un conjunto de pares ordenados, en los cuales las puntuaciones de X ocupan el primer lugar de todos los pares. Al reducir las puntuaciones a las de desviación, X y Y , tendremos otro conjunto de pares ordenados (x, y). Si unas y otras se multiplican y se suman estos productos cruzados, habrá una medida análoga a la suma de dos cuadrados en el análisis de la varianza, que recibe el nombre de *suma de productos cruzados*, o ΣXY . Del mismo modo que el análisis de varianza

(13) No es posible presentar una descripción estadística completa del análisis de covarianza en este libro, cuyo objetivo es familiarizar al estudiante con el método general y mostrarle su relación con otras formas de diseño y análisis estadístico. Pero se puede hallar una exposición buena y sucinta del fundamento de dicho análisis en M. Tate, *Statistics in Education*. Nueva York: Macmillan, 1935, págs. 515-522. En lo que concierne al cálculo es recomendable un libro más antiguo de Edwards: A. Edwards, *Experimental Design in Psychological Research*. Nueva York: Holt, Rinehart and Winston, 1950, cap. 17.

Cuadro 21-2. Análisis del paradigma de covarianza en el estudio de Clark y Walberg

| Grupo experimental (refuerzo masivo) | | Grupo testigo (refuerzo moderado) | |
|---|----------------|--------------------------------------|----------------|
| X (inteligencia) | Y (lectura) | X (inteligencia) | Y (lectura) |
| Medias: 92.05 | 31.62 | 90.73 | 26.86 |

funciona con sumas de cuadrados y varianzas, el de covarianza funciona con las sumas de productos cruzados y covarianzas, lo mismo que con las sumas de cuadrados.

El resultado neto del procedimiento es una tabla de análisis de covarianza, que somete a prueba la significancia de las diferencias de las medias de Y en los grupos experimentales después de ajustar las sumas de Y en los cuadrados. En efecto, el ajuste elimina de dichas sumas la parte resultante de la relación entre X y Y . A una correlación mayor entre ambas corresponde un análisis más eficaz de covarianza. (En caso de que la correlación sea 0 o muy baja, el análisis es una pérdida de tiempo.) Lo que surge para una tabla definitiva del análisis de covarianza son el total intergrupar ajustado y las sumas intragrupalas de los cuadrados. (14) Las varianzas (cuadrados promedio) y la razón F se calculan a partir de esas medidas ajustadas.

El paradigma de un análisis de covarianza se asemeja al diseño de preprueba y posprueba (véase el diseño 19-3). En el cuadro 21-2 aparece un paradigma de la investigación descrita anteriormente

(14) Conviene explicar las tres r que se mencionaron antes. Como una fórmula de r es $r = \Sigma xy / \sqrt{\Sigma x^2 \Sigma y^2}$ y como el análisis de covarianza produce esos términos para el total intergrupar e intragrupal, es evidente que se pueden calcular las tres r . La r para el total es la r usual entre dos series de medidas, y la r de los intragrupos es la correlación de los pares de medias. La contribución clave del análisis de covarianza es que produce la r intragrupal. Puesto que las sumas de dos cuadrados y de los productos cruzados se calculan independientemente en el interior de cada grupo, las diferencias entre ellos no influyen en el cálculo de la r . En consecuencia, la r interna es la "mejor" estimación de la "verdadera" r entre X y Y . Muchas veces pasa desapercibido que algunas r pueden ser aumentadas o disminuidas de manera sorprendente por la varianza intergrupar. Supongamos que una r se calcula entre la inteligencia y aprovechamiento escolar, y que las puntuaciones de ambos sexos figuran en el cómputo. Si hay apreciable diferencia de aprovechamiento entre uno y otro sexo y si la mayor parte de las mujeres se encuentran en un grupo y la mayor parte de los hombres en el otro, esta diferencia intergrupar puede acrecentar demasiado el cálculo de la r . Un análisis de covarianza de los mismos datos arrojará una r intragrupal que probablemente sea menor que la r total original.

donde X = puntuaciones de la prueba de inteligencia y Y = puntuaciones de lectura. (La X también puede representar a la preprueba y Y a la posprueba.) El cuadro es asimismo un paradigma del estudio de Clark y Walberg descrito de manera sucinta en el capítulo 20. Procedemos a examinarlo de nuevo. (15)

Recuérdese que se aplicó a los sujetos del grupo experimental el refuerzo masivo, y el refuerzo moderado a los del grupo de control, y que el análisis simple de varianzas arrojó $F = 9.52$, significativo al nivel de .01. Es posible que las diferencias entre ambos grupos se debiesen a la inteligencia y no al refuerzo. En otras palabras, aunque los sujetos fueron asignados al azar a los grupos experimentales, una diferencia inicial en inteligencia a favor del grupo experimental habría bastado para que la media de su lectura fuese muy superior a la del grupo de control, puesto que la inteligencia guarda estrecha relación con la lectura. En la asignación aleatoria, rara vez sucede esto pero puede ocurrir. Para evitar semejante posibilidad Clark y Walberg se valieron del análisis de covarianza.

Estudie el cuadro 21-2, que presenta un panorama general del diseño y del análisis. Las medias de las puntuaciones de X y Y , registradas por Clark y Walberg, se anotan en la hilera inferior. Las medias de Y son de enorme interés. Aunque seguramente el análisis de covarianza no cambiará este resultado, tal vez la diferencia de las medias de Y (92.05 y 90.73) alteraron las escalas estadísticas a favor del grupo experimental en la prueba de la diferencia entre las medias de Y . El análisis de covarianza de la prueba F , que contiene las sumas de Y en los cuadrados y las medias de cuadrados, libres de la influencia de X , fue apreciable al nivel de .01: $F = 7.90$. En conclusión, hubo gran discrepancia entre los promedios de las puntuaciones de lectura del grupo experimental y del

(15) C. Clark y H. Walberg. "The Influence of Massive Rewards on Reading Achievement in Potential Urban School Dropouts." *American Research Journal*, V (1968), 305-310.

| Métodos | | | | | |
|---------|---|-------|---|-------|---|
| A_1 | | A_2 | | A_3 | |
| X | Y | X | Y | X | Y |

Figura 21-5.

grupo de control, después de ser ajustados o controlados en cuanto a inteligencia.

Pasamos a reseñar un procedimiento discutible y que los expertos han condenado: el empleo del análisis de covarianza en grupos intactos. Lo hacemos porque en numerosas situaciones, especialmente en la investigación pedagógica, es la única vía. El lector no debe olvidar que se trata de una alternativa pobre de la asignación aleatoria.

Una de las principales dificultades de la investigación pedagógica y sociológica es la incapacidad para establecer los grupos experimentales deseados. Es lógico, por ejemplo, que los administradores y maestros se nieguen a fraccionar las clases. Con frecuencia el investigador se ve obligado a mantener intactas las clases y otros grupos. El análisis de covarianza le permite controlar la clase y demás diferencias de grupo mediante técnicas estadísticas. Pongamos un caso: se van a ensayar tres métodos para enseñar ortografía: A_1 , A_2 y A_3 . No es posible usar la asignación aleatoria de los sujetos, pero sí se puede usar el método de clases intactas. Es bien sabido que la inteligencia guarda estrecha relación con la ortografía y que probablemente las clases diferirán mucho en inteligencia. Los métodos se asignan aleatoriamente a clases intactas, y las puntuaciones de la prueba de inteligencia sirven de medidas de X en un análisis de covarianza. El paradigma se muestra en la figura 21-5, donde X = puntuaciones de inteligencia y Y = puntuaciones de ortografía (también las medidas de X pueden representar las puntuaciones de la preprueba de ortografía).

Aunque discutibles, los resultados de este tipo de experimento pueden ser útiles si se les trata con prudencia. El procedimiento tiene ventajas. Una es que las medidas de inteligencia y aprovechamiento normalmente se tienen antes del experimento. De ahí que se les pueda usar sin el peligro de sensibilización que encierran las prepruebas; es posible efectuar un experimento sin que los alumnos se enteren de que se les está aplicando una prueba. La segunda ventaja, íntimamente ligada a la anterior, es la de todo experimento realizado en ambientes naturales, tema que trataremos en la parte 7. La tercera ventaja

consiste en la precisión del análisis y en la información que aporta.

Diseño de Investigación y análisis; observaciones finales

En la preparación y organización de la parte 6 han predominado cuatro objetivos principales. El primero consistió en familiarizar al lector con los principales diseños. Confiábamos que de esa manera lograríamos ensanchar las nociones estrechamente circunscritas a la realización de la investigación, digamos, con un solo grupo experimental y un solo grupo de control, o con sujetos apareados o con un grupo sometido a preprueba y posprueba.

El segundo objetivo tuvo como fin comunicar el sentido de la estructura equilibrada de los buenos diseños. Conviene fomentar una delicada apreciación por la arquitectura del diseño que se ajusta formal y funcionalmente a los problemas de investigación que deseamos resolver.

El tercer objetivo era ayudar al lector a comprender la lógica de la investigación experimental y la de varios diseños. Estas son vías alternativas que llevan al mismo destino: afirmaciones confiables y válidas de las relaciones de las variables. Algunos diseños, si son realizables, producen proposiciones de relación más firmes que otros.

En cierto modo, el cuarto objetivo de la parte 6 fue el más difícil de alcanzar: ayudar a que el estudiante comprendiera la relación existente entre el diseño de investigación y la estadística. Conforme a una de sus acepciones, la estadística es la disciplina que nos da técnicas para manejar las variables. Y hemos visto que uno de los objetivos fundamentales del diseño es controlar las varianzas sistemáticas y de error. Esa es la razón por la cual examinamos a fondo la estadística en las partes 4 y 5 antes de abordar el diseño en la parte 6. Fisher expresa esta idea con precisión al afirmar: "el método estadístico y el diseño experimental no son más que dos aspectos distintos de un mismo todo, que abarca todas las condiciones lógicas del proceso global, el cual consiste en incremen-

tar la ciencia natural por medio de la experimentación". (16)

Un diseño bien ideado no garantiza la validez de los descubrimientos. Los diseños sobrios y bien estructurados para resolver problemas de investigación, pueden arrojar resultados equivocados o conclusiones distorsionadas. Aunque se tienen mayores probabilidades de llegar a conclusiones exactas y válidas, cuando los diseños son adecuados que cuando no lo son, hay algo evidente, si un diseño presenta fallas, se puede llegar a conclusiones confusas. Por ejemplo, si se utiliza un diseño de dos grupos con sujetos igualados, cuando la investigación requiere, de manera lógica, un diseño factorial, o si se emplea un diseño factorial cuando la naturaleza de la investigación indica que se debe usar un diseño de grupos correlacionados, ninguna interpretación o manipulación estadística podrá aumentar la confianza en las conclusiones derivadas de la investigación.

Es justo permitir a Fisher decir la última palabra sobre el tema. En el primer capítulo de su libro, *The Design of Experiments*, señala:

... Si el diseño de un experimento es defectuoso, también lo será cualquier método de interpretación que lo considere decisivo. Es verdad que hay muchísimos procedimientos experimentales que están bien diseñados en cuanto a que pueden conducir a conclusiones decisivas, pero en otras ocasiones sucede así; en tales casos si las conclusiones se extraen a pesar de no estar justificadas, podemos afirmar que la falla se encuentra totalmente en la interpretación y no en el diseño. Pero la falla de la interpretación... consiste en pasar por alto los rasgos distintivos del diseño que hacen que el resultado sea a veces inconclusivo, o concluyente en algunas cuestiones pero no en todas. Comprender bien un aspecto del problema equivale a comprender también el otro. (17)

Guías de estudio

1. ¿Puede mejorarse la memoria con el adiestramiento? William James, el gran filósofo y psicólogo estadounidense, hizo un experimento de memoria consigo mismo. (18) Aprendió 158 versos de un poema de Victor Hugo, tarea para la que necesitó 131½ minutos. Este fue su punto de referencia. Después trabajó unos 20 minutos diarios, durante 32 días, aprendiendo el primer libro de *Paraiso Perdido*. (El libro I consta de 22 páginas compactas de versos bastante difíciles.) De ese modo adiestraba la memoria. Regresó al poema de Victor Hugo y aprendió 158 versos

más en 151½ minutos. Por tanto, necesitó más tiempo después del adiestramiento que antes. No satisfecho con ello, hizo que otros efectuaran trabajos similares y los resultados fueron semejantes.

¿A qué conclusiones pudo haber llegado James a partir de este trabajo? Comente su diseño de investigación. ¿A cuál de los diseños presentados en este libro se asemeja el de James?

2. En un estudio sobre la respuesta que dan los niños rusos y estadounidenses ante la presión de sus compañeros y la de adultos, Bronfenbrenner pidió a 158 estadounidenses de sexto grado y a 188 rusos de quinto grado (edad correspondiente a la de los niños estadounidenses) que respondiesen a situaciones de conflicto con compañeros, adultos y neutrales (anónimas). (19) Entre otras cosas, registró las dos medias que reflejaban la orientación adulta: rusos = 14.82, estadounidenses = 2.22 (a un valor superior correspondía una mayor orientación adulta) y la diferencia era significativa a nivel de .01.

¿Qué problema piensa usted que comportará la interpretación de estos resultados? Concéntrese en los problemas de comparabilidad y control.

3. Kolb llevó a cabo un fascinante experimento con jóvenes de elevada inteligencia que asistían a escuelas de enseñanza media superior y cuyo rendimiento era deficiente, basado en un excelente estudio de McClelland sobre motivación para el aprovechamiento. (20) De 57 jóvenes, asignó 20 aleatoriamente a un programa de adiestramiento en que, con diversos métodos, se les "enseñaba" la motivación para el aprovechamiento, con la intención de crearles la necesidad del logro. Durante el verano se les aplicó una preprueba de motivación para el aprovechamiento y a los seis meses se les volvió a aplicar. En los grupos de control y experimental las puntuaciones medias del cambio fueron respectivamente de 6.72 y .34 significativas al nivel de .005.

a) Comente el uso de las puntuaciones de cambio. ¿Su empleo disminuye nuestra fe en el valor estadístico de los resultados?

b) ¿Es posible que otros factores además del adiestramiento experimental hayan producido el cambio?

4. El diseño factorial de tipo $2 \times 2 \times 2$ es empleado con mucha frecuencia por los psicólogos socia-

(19) U. Bronfenbrenner, "Response to Pressure from Peers Versus Adults among Soviet and American School Children," *International Journal of Psychology*, II (1967), 199-207.

(20) D. Kolb, "Achievement Motivation Training for Underachieving High-School Boys," *Journal of Personality and Social Psychology*, II (1965), 783-792.

les. A continuación citamos dos originales, excelentes y creativos estudios en los que fue utilizado:

Carlsmith, J., y Gross, A. "Some Effects of Guilt on Compliance," *Journal of Personality and Social Psychology*, XI (1969), 232-239.

Jones, E., y col. "Pattern of Performance and Ability Attribution: An Unexpected Primacy Effect," *Journal of Personality and Social Psychology*, X (1968), 317-340.

Lea uno o los dos estudios. (El de Jones es largo, complicado y difícil, pero vale la pena leerlo. Contiene una replicación y una exploración sistemática de hipótesis alternativas y está dotado de gran competencia teórica y técnica.) Fíjese sobre todo en las relaciones de las formulaciones entre los problemas que existen con los diseños.

5. Para que el lector no piense que únicamente se analizan las medidas continuas y que sólo el análisis de varianza se utiliza en los experimentos psicológicos y pedagógicos, lea el estudio de Freedman y col., dedicado al sentido de culpabilidad y a la obediencia. (21) Había un grupo experimental (los sujetos inducidos a mentir) y un grupo de control y la variable dependiente se midió según que un sujeto atendiese o no una petición de ayuda. Los resultados se comunicaron en tablas de frecuencia con divisiones cruzadas.

Lea el estudio y, tras examinar el diseño y los resultados, elabore uno de los tres experimentos en otra forma. Introduzca otra variable independiente; por ejemplo: suponga que se conocía la presencia de amplias diferencias individuales en la obediencia. ¿Cómo se podría controlar eso? Describa y nombre dos clases de diseños con los cuales hacerlo.

(21) J. Freedman, S. Wellington y E. Bless, "Compliance Without Pressure: The Effect of Guilt," *Journal of Personality and Social Psychology*, VII (1967), 117-124.

6. Un medio útil de control mediante el apareamiento consiste en valerse de parejas de gemelos idénticos. ¿Por qué es este método una buena herramienta de control? Si estuviese usted elaborando un experimento para verificar el efecto del ambiente en la medición de la inteligencia y contase con 20 parejas de gemelos idénticos y absoluta libertad experimental, ¿cómo planearía el experimento?

7. En lo que parece ser un estudio clásico sobre la manera de enseñar a los niños a suprimir las fobias (a las serpientes), Bandura y otros utilizaron cuatro grupos experimentales, prepruebas, apareamiento, aleatorización y análisis de la covarianza. (22) Se sugiere a los que hayan realizado estudios avanzados leerlo y explicar el diseño central de la investigación. (Observee que, como en el estudio de Bandura y Menlove citado antes, los sujetos del grupo de control recibieron el tratamiento experimental al terminar el experimento.)

8. En un estudio donde la instrucción relativa a las complejidades de los estímulos artísticos afectaba a la actitud frente a la música entre otras cosas, Renner aplicó el análisis de covarianza. El covariado lo constituían las medidas tomadas de una escala para medir la actitud hacia la música. (23) Era una preprueba. Participaron tres grupos experimentales. Bosquee el diseño basándose en esta breve descripción. ¿Por qué Renner utilizó como preprueba la escala de actitudes ante la música? ¿Por qué uso el análisis de la covarianza? (Nota: vale la pena leer el informe original. El estudio, en parte dedicado a la creatividad, es en sí mismo creativo.)

(22) A. Bandura, E. Blanshard y B. Ritter, "Relative Efficacy of Desensitization and Modeling Approaches for Inducing Behavioral, Affective, and Attitudinal Changes," *Journal of Personality and Social Psychology*, XIII (1969), 173-199.

(23) V. Renner, "Effects of Modification of Cognitive Style on Creative Behavior," *Journal of Personality and Social Psychology*, XIV (1970), 257-262.

(16) R. Fisher, *The Design of Experiments*, 6a. ed. Nueva York: Hafner, 1951, pág. 3.

(17) *Ibid.*, págs. 2-3.

(18) W. James, *The Principles of Psychology*. Nueva York: Holt, 1890, págs. 666-667.