

PARTE II
METODOLOGÍA CUANTITATIVA

CAPÍTULO 5

MÉTODOS DE INVESTIGACIÓN DE ENFOQUE EXPERIMENTAL

Antoni Sans Martín

OBJETIVOS

Con el estudio de este capítulo se podrá:

1. Criticar de modo reflexivo y fundamentado informes de investigación de enfoque experimental publicados en revistas científicas especializadas.
2. Identificar los fundamentos en las que se basa esta perspectiva epistemológica.
3. Formular y reconocer las fases fundamentales de una investigación de enfoque cuantitativo.
4. Conocer e identificar los procedimientos más importantes para controlar las variables extrañas en un diseño experimental.
5. Distinguir entre diseños preexperimentales, experimentales y cuasiexperimentales.

INTRODUCCIÓN HISTÓRICA

Desde la perspectiva histórica, Buyse (1949: 592) indica tres influencias iniciales en el método experimental: el pensamiento filosófico dominante en el siglo XIX; la evolución de la psicología tradicional, que provocó el nacimiento y el desarrollo del espíritu científico; y el desarrollo del método experimental. Durante el siglo XIX predominaban corrientes filosóficas como el positivismo (A. Comte, 1798-1857), el sociologismo (E. Durkheim, 1858-1917), el pragmatismo (W. James, 1840-1910) y el experimentalismo (J. Dewey, 1859-1952). Estas concepciones filosóficas fueron decisivas para la emancipación de las ciencias y tuvieron repercusiones importantes en el desarrollo del método experimental.

El positivismo surgió bajo la influencia de sistemas como el empirismo británico, tanto de Locke y de Hume como de la corriente historicista de Vico; su representante más significativo fue Comte. Según esta concepción no hay más conocimiento válido que el positivo, caracterizado por:

- La explicación de los fenómenos a partir de ciertas regularidades o leyes establecidas por la observación de fenómenos observables.
- La utilización de la observación sistemática, la experimentación y la comparación como estrategias fundamentales para generar el conocimiento.
- La capacidad predictiva de los enunciados científicos.
- La comprobación empírica de los enunciados científicos, que se alcanzará con un control empírico similar al que se usa en el caso de los fenómenos naturales.
- La finalidad tecnicoutilitaria del conocimiento; lo importante es la vertiente aplicada: la práctica garantiza la teoría.

La epistemología sociológica de Durkheim intentó sintetizar positivismo e idealismo, criticando las insuficiencias del método positivista aplicado a las ciencias sociales. Este autor, fascinado por el método de C. Bernard, intentó dar rigor a las ciencias sociales y asignó a la sociología:

- 1) Un *objeto*: hechos sociales, como por ejemplo las características morfológicas y demográficas, las normas legales e institucionales, las creencias y prácticas establecidas. Los hechos sociales, según esta concepción, tienen tres características: a) son externos a los individuos, anteriores y superiores a las ciencias individuales; b) son generales en una sociedad e independientes de sus manifestaciones individuales; c) se imponen coercitivamente sobre el individuo, justifican el objetivismo metodológico y en definitiva la universalidad del método.
- 2) Un *método*: a estos hechos sociales se les puede aplicar un método positivo en la medida en que son cosas o datos reales y observables: constituyen el punto de partida de la ciencia.

W. James adaptó las concepciones del *pragmatismo* fundado por Ch.S. Peirce el año 1878 y volvió a poner énfasis en las cosas

concretas, los hechos y la acción, con predominio de la orientación empirista. La práctica era un valor absoluto y se constituía en el criterio final de toda formulación científica, de manera que se llegaba a identificar la verdad con la utilidad. Así, el valor de las teorías dependía de su viabilidad para resolver problemas inmediatos. Según esta corriente la realidad se va haciendo continuamente, y parte de su estructura pertenece al futuro, mientras que según el racionalismo la realidad ya estaba hecha y completa (determinismo).

J. Dewey (1859-1952) aplicó el pragmatismo de W. James a las ciencias sociales, y dio origen así al *instrumentalismo* o *experimentalismo*, que se basaba en la experiencia y en la práctica. Estas concepciones y estos planteamientos filosóficos predominantes en el siglo XIX influyeron en la manera de concebir la naturaleza de los fenómenos educativos y contribuyeron bastante a configurar el carácter científico de la educación.

Las concepciones metodológicas serán desarrolladas sobre todo por los hombres de ciencia comprometidos en la actividad investigadora, como por ejemplo Pascal, Newton y Claude Bernard; y al mismo tiempo por teóricos puros, que se dedicarán al análisis del método experimental, como por ejemplo Whewell y Stuart Mill. La metodología experimental alcanza su punto culminante en la obra de Claude Bernard (1865), que expone claramente los principios esenciales del método de experimentación:

- 1) El científico se somete a los hechos, a los cuales tiene que sacrificar la teoría por muy brillante que sea.
- 2) La investigación experimental se basa en el determinismo de los fenómenos, que se tendrán que repetir en las mismas condiciones hasta que entre ellos se puedan establecer relaciones constantes.
- 3) La ciencia es ajena a cuestiones epistemológicas, puesto que no se pueden comprobar experimentalmente.
- 4) Las hipótesis se contrastan decisivamente con las contrapruebas.

Fisher (1890-1962) dio un paso importante con *The Design of Experiments* (1935), considerada como uno de los referentes básicos en el diseño experimental basado en el análisis de la varianza.

La obra de Campbell y Stanley (1966) representó un avance significativo en la evolución de las técnicas experimentales. Como

un paso más de esta maduración metodológica había una tendencia a explorar nuevos diseños que se ajustasen más bien a las condiciones naturales; por ejemplo, diseños cuasiexperimentales (Cook y Campbell, 1979) y de sujeto único (Kratowill, 1978).

Otra tendencia introducía la orientación cualitativa, como consecuencia de la incidencia de la etnografía y la sociología en la investigación educativa. Esta situación provocó controversias metodológicas entre estas posiciones (cuantitativa vs. cualitativa) que fueron objeto de numerosos artículos de revistas especializadas.

La introducción del ordenador en la investigación representó una gran revolución metodológica, que permitió desarrollos estadísticos y técnicos –métodos multivariantes– y provocó notables cambios en el enfoque y el planteo de la investigación nomotética. La disponibilidad de programas permitió no sólo hacer análisis complejos de datos, sino también formar bases de datos que permitían generalizar estudios bibliométricos y de metanálisis.

CARACTERÍSTICAS DE LA INVESTIGACIÓN EXPERIMENTAL

La investigación experimental tiene seis características distintas:

TABLA 5.1. Características de la investigación experimental

<ol style="list-style-type: none"> 1. Equivalencia estadística de sujetos en diversos grupos normalmente formados al azar 2. Comparación de dos o más grupos o conjuntos de condiciones 3. Manipulación directa de una variable independiente 4. Medición de cada variable dependiente 5. Uso de estadísticos inferenciales 6. Diseño que permita un control máximo de variables extrañas

1) La primera característica de la investigación experimental es la reunión de sujetos en **grupos equivalentes**. Por tanto, ninguna de las diferencias de los resultados puede ser atribuida a diferencias iniciales entre los grupos. El método usado más a menudo para formar los grupos es la asignación al azar de los sujetos.

- 2) La segunda característica es la necesidad de que haya **dos grupos** como mínimo para establecer comparaciones. Un experimento no se puede llevar a cabo con un grupo de sujetos y una única condición experimental. Un experimento implica comparar el efecto de una condición sobre un grupo con el efecto de una condición sobre otro grupo, o comparar el efecto de condiciones diferentes en el mismo grupo. Estos dos grupos o condiciones son necesarias para establecer la comparación.
- 3) La manipulación de **variables independientes** es, tal vez, la característica más distintiva del enfoque experimental. *Manipulación*, en este ámbito, quiere decir que es el investigador quien decide sobre cada uno de los niveles que corresponderá a cada grupo de sujetos. La variable independiente o causa se manipula en forma de diferentes valores o condiciones (niveles) que el experimentador asigna. Si las condiciones no las asigna el investigador, el estudio no puede considerarse un experimento real.
- 4) La cuarta característica de la investigación experimental, la medición de **variables dependientes**, hace referencia a fenómenos que pueden ser consignados con valores numéricos. Si el resultado del estudio no puede ser medido ni cuantificado de este modo, la investigación difícilmente podrá ser experimental.
- 5) Otra característica que comporta valores numéricos es la utilización de **estadística inferencial**. La estadística se usa para tomar decisiones en términos de probabilidad, lo cual nos permite hacer generalizaciones a partir de muestras de sujetos.
- 6) La última característica es el **control de variables extrañas**. Controlamos variables extrañas para asegurarnos que estas variables no influyen en la variable dependiente, o si influyen, lo hacen de modo homogéneo en todos los grupos.

Dichas características son habituales en la investigación experimental en medicina y biología, pero en la investigación pedagógica muy raramente se pueden seguir en su totalidad. Pero eso no hace que disminuya la importancia del método experimental en educación. Muchas investigaciones se acercan a las características de un experimento puro, y por esta razón es preciso aplicar de modo diferente métodos de investigación del comportamiento.

to que identifiquen las causas que permitan la interpretación de los fenómenos. Esta era la idea de Campbell y Stanley (1963) cuando escribieron su conocido capítulo “Experimentación y diseños cuasiexperimentales para la educación”.

FASES DE UN EXPERIMENTO

De hecho, un experimento consiste en una situación simulada en la que se modifican voluntariamente las condiciones de una o de diversas situaciones precedentes –variable independiente– para comprobar como afecta esta variable independiente otra situación consiguiente –variable dependiente– que se observa sistemáticamente. Esta situación se plantea en diversas etapas, que pueden resumirse tal y como se presenta a continuación.

Planteamiento de un problema de conocimiento

Cuando no tenemos una explicación satisfactoria para un fenómeno o tenemos una laguna de conocimiento se nos plantea un problema de conocimiento. El problema está relacionado con la variable dependiente. Así, un investigador que se dedica a los métodos de estudio, se hará preguntas sobre la naturaleza de estos métodos, sus efectos, los procesos cognitivos que están asociados, etc. Las preguntas que se planteen tienen que poder resolverse con la aplicación de un procedimiento experimental, lo cual sólo es posible en determinados casos. La naturaleza del problema determinará en gran parte el enfoque metodológico.

Formulación de hipótesis

A partir de un problema determinado, un investigador habrá de **anticipar** un resultado posible de la investigación experimental. Esta anticipación es la **hipótesis**. O sea, la hipótesis es la relación apriorística entre la variable independiente o causa y la variable dependiente o efecto. La hipótesis, en cualquier caso, se deberá comprobar empíricamente.

Realización de un diseño adecuado a la hipótesis

El diseño de un experimento, tal y como veremos más adelante, es su organización formal e incluye diversos subprocesos,

como por ejemplo la definición de las variables independientes con sus niveles correspondientes, la asignación de los sujetos a los diferentes grupos experimentales, la observación sistemática de resultados, etc.

Recogida y análisis de datos

Para recoger datos se pueden usar diferentes instrumentos y técnicas que pueden ser desde impresos (tests, instrucciones escritas, registros de observación, etc.) hasta aparatos complejos (temporizador, osciloscopio, taquiscopio, etc.). Esto implica tomar en consideración las características de la medición educativa; para detalles sobre la problemática de la medición educativa véase Mateo (2004) en esta misma colección.

Una vez obtenidos los resultados, se han de analizar siguiendo un plan de análisis de datos que se habrá de haber previsto antes, para evitar situaciones irresolubles. Los procedimientos más frecuentes son los de carácter estadístico, aún cuando es posible usar modelos lógicos o modelos matemáticos no numéricos. Para detalles sobre el análisis de datos, véase Etxeberria y Tejedor (2004) y Tejedor y Etxeberria (2004) en esta misma colección.

Elaboración de conclusiones

En las conclusiones se especifica hasta qué **punto** y con **qué limitaciones** queda confirmada la hipótesis experimental, y también el posible **grado de generalización** que nos permite, con arreglo al proceso que se ha seguido y con la muestra con el que se ha trabajado. Unas buenas conclusiones suelen incluir problemas para resolver que han surgido a lo largo del proceso investigador. A veces, se puede valorar un experimento más por las perspectivas que abre que por sus resultados específicos.

ESTUDIO DE LAS VARIABLES

Tal y como hemos podido comprobar, el enfoque experimental usa un lenguaje formalizado que en términos generales coincide con el lenguaje científico. Este sistema de conceptos aplicable a hechos se formaliza con un lenguaje creado expresamente para explicar una parcela de la realidad. En palabras de Kerlinger

(1981), un **concepto** es una abstracción formada por una generalización de casos particulares. Dicho de otra manera, los conceptos se forman en procesos de observación y reflexión a partir de los acontecimientos o los objetos. Un **concepto científico** o **constructo** es aquél que se crea deliberadamente para explicar la realidad. La creación de un constructo requiere una definición o especificación, y también que se le incardine en una teoría en la que se relacione con otros constructos.

Mientras que los conceptos se enmarcan en el ámbito del conocimiento vulgar, los constructos lo hacen en el contexto de sistemas científicos. En los procedimientos de investigación de enfoque experimental, se usan determinados constructos que tienen la propiedad de poder ser medidos; este tipo de constructos son los generalmente denominados variables.

La medición de una variable es un proceso con el cual se representan propiedades por medio de números, siguiendo unas reglas determinadas. Es decir, se pasa de un constructo, que se puede explicar conceptualmente (por ejemplo inteligencia), a una variable que puede ser medida (el test de Binet). A esto se le llama la operativización de una variable, que consiste en definir una variable de forma operativa mediante las operaciones que se siguen para su medición o manipulación.

Desde la perspectiva de uso del criterio metodológico, vista su naturaleza, las variables se pueden clasificar como: variables independientes, variables extrañas, variables contaminantes, variables controladas y variables dependientes.

Variables independientes

La variable independiente es la supuesta causa de los cambios observados al término del experimento en la variable dependiente. Por definición, la variable independiente se debe poder manipular; o, en su defecto puede ser asignada por el investigador. Es el antecedente de la predicción establecida en la hipótesis del proceso de investigación.

Variables dependientes

La variable dependiente es la que recoge los efectos producidos por la variable independiente. Es la variable que está relacio-

nada con el problema investigado. Es preciso aclarar que la variable dependiente y la independiente lo son con arreglo a la hipótesis del experimento, de manera que se podría dar el supuesto de que una variable que sea dependiente en un experimento sea independiente en otro. Supongamos, por ejemplo, que queremos conocer el efecto que produce un programa de mejora de la autoestima sobre el rendimiento académico. En tal caso, la autoestima es la variable independiente y el rendimiento, la variable dependiente. En otro planteamiento podríamos querer conocer qué efectos produce el rendimiento académico sobre la autoestima del sujeto. En este último caso, la variable independiente sería el rendimiento y la variable dependiente, la autoestima.

Variables extrañas

Por exclusión, podemos definir las variables extrañas como las que no son ni la variable independiente ni la dependiente. En cualquier planteamiento investigador hay variables, además de la independiente, que afectan las condiciones del experimento. Es preciso suponer que debido al azar los valores de estas variables se distribuyen homogéneamente entre los diferentes grupos experimentales.

Variables contaminantes

Consideremos variable contaminante la que, sin que el investigador lo quiera, añade sus efectos a los de la variable independiente y falsea así los resultados finales que se recogen en la variable dependiente. Para que una variable se pueda considerar contaminante, es preciso que añada sus efectos sistemáticamente sólo a algún grupo del experimento. Hay que vigilar especialmente la influencia de este tipo de variables, ya que a menudo pueden invalidar totalmente o en parte las conclusiones de una investigación.

Variables controladas

Son aquéllas que se igualan entre los diferentes grupos que se forman en un experimento para evitar que influyan de forma desigual. Al efecto de controlar variables debe tenerse en cuenta que las diferentes condiciones o niveles de la variable indepen-

diente estén correctamente aplicadas, de manera que el tratamiento que reciba cada grupo sea sólo el previsto en la hipótesis. Es preciso controlar todas las variables que puedan ser contaminantes y la homogeneidad de las cuales no podamos tener garantizada con la formación de los grupos al azar.

CONTROL DE VARIABLES EXTRAÑAS

Un aspecto esencial del diseño de un experimento consiste en asegurarnos que los resultados no sean interferidos por los efectos de variables extrañas a la variable independiente. O sea, se trata de igualar los grupos por lo que se refiere la variancia secundaria. La variancia primaria es la debida a la variable independiente; la variancia secundaria es la debida a las variables extrañas, que conviene controlar. Acto seguido describiremos algunos de los métodos de control más usuales.

Aleatorización

Se trata del procedimiento más importante y más utilizado para controlar las variables extrañas. Consiste en elegir al azar o aleatoriamente (sujetos, condiciones, grupos, centros, aulas, profesores, etc.). Tiene una virtud carente en los otros: teóricamente es el único que sirve para controlar **variables desconocidas**.

Siguiendo a Cochran y Cox (1978), citados por Rosell (1986): "la aleatorización es, en cierto modo, como un tipo de seguro, por el hecho de que es una precaución contra inferencias que pueden o no ocurrir, y ser o no ser importantes si ocurren". Generalmente es aconsejable utilizar la aleatorización como procedimiento por defecto; es decir, hacerlo sistemáticamente en el momento de elegir sujetos, grupos, condiciones experimentales, formar parejas y en cualquier situación del desarrollo de un experimento.

Mantenimiento constante de variables extrañas

En determinadas ocasiones, cuando se quiere controlar una variable supuestamente contaminante, será aconsejable mantenerla a un nivel medio constante, de manera que las condiciones en las que realicemos la investigación se asemejen tanto como sea posible a las reales.

Emparejamiento de los sujetos

Cuando sospechemos que los grupos que se tienen que formar pueden diferir en alguna característica que pueda afectar decisivamente los resultados, una posibilidad de control es el emparejamiento. Se trata de formar parejas que se asemejen en el valor de esta característica o variable de emparejamiento que, en cualquier caso, ha de estar estrechamente relacionada con la variable dependiente. En un estudio sobre el rendimiento académico podrían formarse parejas de alumnos que compartan determinada dificultad de aprendizaje para posteriormente asignar al azar un sujeto al grupo experimental y el otro al grupo control. Este tipo de control aumenta cuando se dispone de hermanos gemelos.

Bloqueo

A veces los grupos que es preciso formar incluyen más de dos sujetos que constituyen un bloque homogéneo con arreglo a una o más variables de formación de bloques. Una vez hechos los bloques, los sujetos son asignados al azar a cada grupo de tratamiento. La formación de bloques es una generalización del procedimiento de emparejamiento. Hay procedimientos estadísticos especialmente indicados para analizar la variancia en tal caso.

Balanceo

Con este procedimiento distribuimos del mismo modo la presencia de una variable extraña en los grupos que formemos. En primer lugar se identifican las variables extrañas que puedan influir en los resultados del experimento (experimentadores, características de los sujetos o procedimientos de obtención de datos). Acto seguido se forman los grupos igualando las variables extrañas. Imaginemos que en un experimento intervienen tres profesores (A, B, C), alumnos de ambos sexos y dos condiciones experimentales (método tradicional y método activo). El número de condiciones posibles en tal caso será $3 \times 2 \times 2 = 12$. O sea, se deben hacer las variaciones sin repetición de los elementos de los diferentes grupos. Suponiendo que la muestra estuviese formada por 120 sujetos y que existiera el mismo número de alumnos de cada sexo, el montante de sujetos para cada combinación posible

sería $120 / 12 = 10$ sujetos por grupo. La combinación de condiciones experimentales está equilibrada respecto de cada variable experimental. La figura 5.1. ilustra el diseño balanceado para esta investigación.

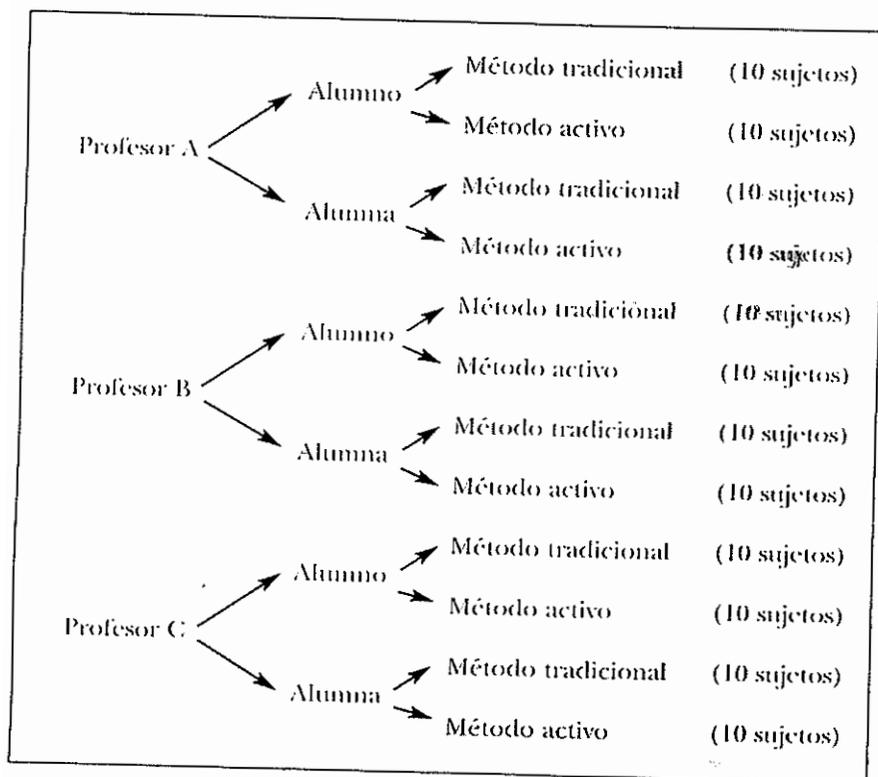


Fig. 5.1. Diseño balanceado

Contrabalanceo

Este procedimiento de control de variables extrañas se utiliza en situaciones experimentales **intrasujeto**, cuando un mismo sujeto debe pasar por diversos procedimientos experimentales. En tal caso interesa controlar los efectos de tipo secuencial o de orden que se puedan producir. La influencia de la fatiga y del

aprendizaje que se puede crear con la aplicación sucesiva de diversos tratamientos se suele denominar "error progresivo". Precisamente el contrabalanceo es el procedimiento usual para neutralizarlo. En tal caso, el número de condiciones experimentales es $n!$. La permutación de n elementos, siendo n el número de tratamientos experimentales. Recordemos que la permutación de n elementos se simboliza por $n!$ y su cálculo viene dado por: $n! = 1 \times 2 \times \dots \times n$. Por ejemplo, si queremos comprobar la eficacia de tres Métodos (M1, M2 y M3) y hemos optado por una situación intrasujeto, podemos organizar $3! = 1 \times 2 \times 3 = 6$ condiciones experimentales. Si dispusiéramos de 60 sujetos, los que pasarían por cada una de las condiciones anteriores serían 10 ($60 : 6 = 10$). Podemos representar las diferentes condiciones de la manera siguiente:

- Condición 1: M1 M2 M3 (10 sujetos)
- Condición 2: M1 M3 M2 (10 sujetos)
- Condición 3: M2 M1 M3 (10 sujetos)
- Condición 4: M2 M3 M1 (10 sujetos)
- Condición 5: M3 M1 M2 (10 sujetos)
- Condición 6: M3 M2 M1 (10 sujetos)

Cuando el número de condiciones experimentales es muy alto, dado que necesitamos que haya igual o mayor número de sujetos, a veces se hace difícil disponer de muestras suficientemente grandes. En tal caso, se puede hacer un contrabalanceo incompleto.

Control estadístico

Todos los procedimientos expuestos hasta ahora intervienen en la formación de grupos o en la administración de las condiciones experimentales. En ciertas ocasiones no es posible este control. Estas situaciones son habituales, bien porque los grupos ya están formados, bien porque difieren en lo que se refiere a alguna variable muy correlacionada con la variable dependiente. El procedimiento más usual de control estadístico es el ANCOVA (ANalysis of COVariance; análisis de la covarianza). Esta técnica permite comparar los diferentes grupos y eliminar la influencia de la covariable (variable extraña relacionada con la variable dependiente). Supongamos que la inteligencia está correlaciona-

da con la autoestima y que para experimentar un programa de aumento de la autoestima no podemos aplicar un procedimiento de control experimental, puesto que tenemos que trabajar con grupos ya formados. Podríamos controlar la variable inteligencia (covariable) con un control estadístico a partir de la técnica del ANCOVA, que nos permite comparar los diferentes grupos experimentales "ajustando" los cambios observados en la variable dependiente eliminando la influencia de la inteligencia. Si los grupos son muy grandes, el control estadístico tiende a ser tan seguro como el experimental.

LIMITACIONES DE LA INVESTIGACIÓN DE ENFOQUE EXPERIMENTAL

La realidad es enormemente compleja, cambiante y diversa. Cuando la estudiamos nos encontramos con limitaciones y obstáculos, algunos de los cuales son difíciles de superar.

Limitaciones de tipo ambiental

Las limitaciones de este tipo se refieren a las situaciones ambientales (condiciones del ambiente, características de los sujetos, etc.) que pueden afectar los resultados de la investigación. En la investigación de enfoque experimental, las variables que intervienen y que difieren de unos contextos a otros (edad, sexo, nivel sociocultural, etc.) son múltiples, lo cual hace que los resultados sean aplicables al ámbito estudiado, pero a veces son difíciles de generalizar a otros ámbitos. La capacidad de un diseño para generalizar resultados generalmente se denomina "validez externa". Las condiciones ambientales determinan el proceso de generalización y ponen límites al alcance de los resultados de investigación.

Limitaciones en la medida

La medida presenta dificultades de observación y cuantificación, atendidas las características de los fenómenos humanos, más complejos que los naturales. El mundo psíquico o interior de los sujetos (significados, motivaciones, etc.) es difícil de medir, porque no es observable directamente y es preciso acceder a partir de sus manifestaciones. Los instrumentos y las técnicas de

recogida de datos disponibles (tests, cuestionarios, sistemas de categorías, etc.) no tienen el grado de precisión y exactitud de los instrumentos utilizados en otras ciencias (rayos X, microscopio, etc.) y eso hace más difícil llegar a conocer la realidad. En los fenómenos psicológicos o internos se suelen aplicar medidas de carácter ordinal y juicios de estimación; sin embargo, los supuestos implícitos en las escalas de medida son cuestionables. Cuanto más potentes y precisos, o sea, válidos y fiables, sean los instrumentos de medida que tengamos a nuestro alcance, esta limitación disminuirá.

Limitaciones de tipo moral

La investigación con seres humanos está limitada por condiciones de tipo moral que afectan a los sujetos que participan. Las limitaciones morales se refieren a aspectos que influyen claramente en las personas, siempre y cuando sea de modo perjudicial (en la personalidad, en el desarrollo físico, emocional, en la intimidad, etc.). La investigación, para que sea moralmente lícita, debe respetar los derechos inalienables de la persona humana, tal y como recogen los códigos éticos dictados por asociaciones profesionales.

Limitaciones derivadas del objeto

La naturaleza de la realidad educativa hace difícil que la podamos conocer bien. El problema que se plantea es si la investigación ha de ocuparse sólo de la realidad observable y cuantificable o también tiene que penetrar en la no observable (significados, intenciones, creencias, etc.). Para el neopositivismo, el conocimiento válido es el que se obtiene con las ciencias experimentales, que se limitan a relacionar los fenómenos o los datos positivos sin pretender ir más allá del mundo empírico. Las ciencias quedan reducidas entonces a aquello que es verificable empíricamente, y fuera de este ámbito no existe ningún otro conocimiento válido de los hechos. La fiabilidad de las ciencias positivas no es suficiente ni se puede extender a toda la realidad. Desde el positivismo se puede conseguir fiabilidad sólo cuando se estudian aspectos de la realidad que son materiales, que están sometidos a leyes constantes, que se pueden someter al control propio del

método experimental; muchos problemas humanos no se pueden resolver desde el positivismo; son problemas que afectan a las personas en su ser más profundo. La ciencia no es una panacea (Bunge, 1976: 49) y la tenemos que considerar la vía más adecuada para resolver algunos problemas o para dar respuesta a algunos interrogantes; pero para otros problemas o interrogantes, la ciencia no tiene respuesta.

DISEÑOS MÁS CORRIENTES DE ENFOQUE EXPERIMENTAL EN INVESTIGACIÓN EDUCATIVA

Con la finalidad de analizar más sistemáticamente los diseños que se expondrán a continuación, exponemos las principales fuentes que invalidan internamente un diseño para que nos sirvan de guía para su clasificación y análisis.

<i>Fuentes de invalidez interna de un diseño experimental</i>	
ASPECTO	DESCRIPCIÓN
<i>Historia</i>	Son los acontecimientos no previstos o extraños que afectan al resultado, que pueden ocurrir mientras se lleva a cabo la investigación. Si la investigación dura mucho tiempo es probable que se produzca este efecto.
<i>Selección</i>	Se trata de diferencias entre los sujetos de los grupos, que pueden ser debidas a su composición. Un muestreo sesgado puede producir un efecto de selección.
<i>Regresión estadística</i>	Regresión: tendencia de los valores extremos a "regresar" a los valores medios. El efecto regresión consiste en que si en el pretest se dan valores muy altos o muy bajos, en el posttest probablemente se produzca una "regresión" a los valores medios, independientemente de los efectos de las VI.
<i>Testing</i>	Ocurre cuando, al aplicar un test o cuestionario, éste afecta a los sujetos produciendo resultados superiores solo por el hecho de haberles aplicado un test previamente.

<i>Instrumentación</i>	Son las diferencias en los resultados debidas a la falta de fiabilidad de los instrumentos de medida o de los observadores, entre el pretest y el posttest.
<i>Difusión del tratamiento</i>	Ocurre cuando los sujetos de un grupo aprenden algo sobre el tratamiento o las condiciones de otro grupo. Por ejemplo, cuando el grupo control "aprende" a partir del grupo experimental, debido a una difusión de la información.
<i>Influencia del experimentador</i>	Se produce cuando existen efectos, intencionados o no, producidos por el investigador, en las respuestas de los sujetos.
<i>Condiciones estadísticas</i>	Es la violación de las condiciones de aplicación o selección incorrecta de la prueba estadística
<i>Mortalidad</i>	Ocurre cuando se produce una pérdida sistemática (no aleatoria) de los sujetos experimentales o del grupo control.
<i>Maduración</i>	Ocurre cuando algún efecto se debe a la maduración de los sujetos y no al aprendizaje a partir del programa.

Las características descritas en esta relación servirán para analizar las posibilidades y las limitaciones de cada tipo de diseño. En la presentación de los diseños se utilizará el sistema de representación universal, de modo similar a la notación que usan Cook y Campbell (1979) y Campbell y Stanley (1963). La significación de la notación es la siguiente

- R: Aleatorización (R del inglés *random*, 'azar').
- O: Observación, medida registrada en el pretest o en el posttest.
- X: Tratamiento (los subíndices 1 a n indican diferentes tratamientos).

DISEÑOS PREEXPERIMENTALES

A los tres diseños que se describen a continuación, denominados preexperimentales, les faltan dos o más de las seis características con que se han definido las características de la investiga-

ción experimental (Tabla 5.1.). En consecuencia, algunas características de su validez interna son muy débiles y por ello los resultados son difíciles de interpretar. Por esta razón no tienen suficiente fuerza para hacer inferencias causales. Sin embargo, se suelen usar como procedimientos para generar ideas que, más adelante, puedan ser probadas con diseños más sistemáticos.

Diseño de sólo posttest con un grupo

En este diseño el investigador proporciona un tratamiento y a continuación hace una observación (es medida la variable dependiente); eso se puede representar en el diagrama siguiente:

Grupo	Asignación	Pretest	Tratamiento	Posttest
A	no R		X	O

En este diseño no se pueden aplicar todas las características de validez interna, porque no hay pretest y no se considera la comparación con otros tratamientos. Tan solo podemos hacer aproximaciones por lo que respecta a las relaciones causales. Sin pretest resulta difícil concluir que la conducta haya cambiado en algo. Sin parangón o grupo de control también es difícil conocer la influencia de otros factores acaecidos en el mismo tiempo que se ha producido el tratamiento y que pueden haber influido en los valores de la variable dependiente. Por ejemplo, imaginemos un profesor de métodos de investigación que quiere hacer un estudio sobre los resultados obtenidos por los alumnos en sus clases. Parece razonable suponer que no conocen prácticamente nada del contenido del curso antes de iniciarlo, y en consecuencia podemos deducir que la observación del posttest se corresponde sólo al tratamiento (clases).

Diseño de pretest-posttest con un grupo

Este diseño sólo se distingue del anterior en el pretest que precede el tratamiento. Se puede representar según el diagrama siguiente:

Grupo	Asignación	Pretest	Tratamiento	Posttest
A	no R	O	X	O

En este diseño se aplica un pretest (O) a un grupo de sujetos, después el tratamiento (X) y finalmente el posttest (O). El resultado es la valoración del cambio ocurrido desde el pretest hasta el posttest. De hecho este diseño se ha popularizado como diseño pretest-posttest. Mientras que el investigador puede obtener una medida del cambio, sin embargo no puede comprobar hipótesis alternativas.

Diseño de sólo posttest con dos grupos no equivalentes

Este diseño se asemeja al diseño de sólo posttest en un grupo. La diferencia es que en este caso hay un grupo control que no recibe el tratamiento. Este diseño puede representarse con el diagrama siguiente:

Grupo	Asignación	Pretest	Tratamiento	Posttest
A	no R		X	O
A	no R			O

A menudo se utiliza este diseño antes de implementar un tratamiento. La expresión "grupos no equivalentes" se usa porque la selección de los sujetos no se produce en condiciones aleatorias que garanticen totalmente la validez interna. Las diferencias iniciales entre los grupos de sujetos pueden influir en los resultados del posttest.

DISEÑOS PROPIAMENTE EXPERIMENTALES

En este apartado presentamos dos diseños que se pueden denominar diseños experimentales verdaderos; incluyen procedimientos para extraer diferencias intersujeto por medio de la formación de grupos al azar y la manipulación de la variable de tratamiento o variable independiente. Son los diseños que históricamente se han llamado "experimentales" en las ciencias físicas y biológicas.

Diseño de pretest-posttest con grupo de control

Es una extensión del diseño pretest-posttest con un grupo. En tal caso se añade al grupo experimental (grupo A) un segundo grupo denominado "grupo de control" (grupo B) o "de compa-

ración”, y los sujetos se asignan a cada grupo al azar. Se puede representar con el diagrama siguiente:

Grupo	Asignación	Pretest	Tratamiento	Posttest
A	R		X	O
A	R			O

- 1) El **primer paso** consiste en asignar los sujetos al grupo experimental (o de tratamiento) y al grupo de control, al azar. En estudios con un volumen relativamente pequeño de sujetos es preferible ordenarlos a partir de calificaciones, actitudes u otros factores que estén relacionados con la variable dependiente. En el caso de diseños de dos grupos, se pueden formar parejas de sujetos de características similares (emparejamiento de sujetos). El investigador asigna un sujeto a cada grupo. Otro procedimiento consiste en incluir sujetos con arreglo a una variable y entonces asignar aleatoriamente cada miembro al grupo experimental y al grupo control. En el caso de grupos pequeños conviene comprobar que no haya diferencias iniciales. Si los grupos son numerosos, por ejemplo de 200 personas o más, no hemos de esperar grandes diferencias entre los grupos si se han formado al azar. En general, los investigadores suelen trabajar con grupos de 30 sujetos o más cada uno, con la finalidad de asumir los supuestos estadísticos paramétricos. Estos supuestos tienden a cumplirse con muestras grandes ($n > 30$).
- 2) El **segundo paso** consiste en aplicar a cada grupo un pretest sobre la variable dependiente (VD).
- 3) El **tercer paso** consiste en administrar el tratamiento (variable independiente, VI) sólo al grupo experimental, procurando que todas las otras condiciones sean equivalentes en los dos grupos, de manera que la única diferencia sea la influencia del tratamiento. A esto se le denomina *caeteris paribus* (todas las demás condiciones equivalentes).
- 4) En el **cuarto paso**, cada grupo es sometido a un posttest sobre la variable dependiente. Finalmente se analizan los resultados.

Diseño sólo de posttest con grupo control

La utilización del azar para formar el grupo experimental y de control los iguala antes de introducir la variable independiente. Si los grupos se han igualado por medio del azar, ¿es preciso aplicarles un pretest? Si los grupos se componen como mínimo de 15 individuos cada uno, no es preciso; en tal caso no es indispensable aplicar un pretest para llevar a cabo un estudio experimental. La única diferencia con el diseño de pretest-posttest con grupo control es precisamente la ausencia del pretest. Este diseño se puede representar con el diagrama siguiente:

Grupo	Asignación	Pretest	Tratamiento	Posttest
A	R		X	O
B	R			O

Usaremos este diseño cuando no sea posible o conveniente hacer un pretest, o en situaciones en las que el pretest pueda influir directamente en el efecto del tratamiento.

Diseño Solomon

Para satisfacer las necesidades de control por una parte y la inconveniencia de aplicar un pretest por otra, Solomon ideó un diseño que lleva su nombre. Tiene cuatro grupos, dos experimentales y dos de control. El pretest se pasa a dos grupos (uno experimental y otro de control); mientras que a otros dos grupos (uno experimental y otro de control) solo se les administra el posttest. De esta forma se puede controlar los efectos del pretest sobre el posttest. El diagrama siguiente ilustra el diseño Solomon.

Grupo	Asignación	Pretest	Tratamiento	Posttest
A (experimental)	R	O	X	O
B (control)	R	O		O
C (experimental)	R		X	O
D (control)	R			O

El diseño Solomon reúne una serie de condiciones que lo convierten en uno de los más recomendables siempre que sea

posible. Una variación de este diseño es la no asignación al azar de los sujetos; lo cual se puede hacer sin afectar gravemente a los resultados cuando los grupos son grandes y se asegura el mantenimiento de las condiciones *caeteris paribus*.

DISEÑOS CUASIEXPERIMENTALES

Los diseños experimentales proporcionan los argumentos más sólidos y convincentes para establecer efectos causales de la variable independiente, por el hecho de que se controlan bien las fuentes de validez interna. A menudo se dan circunstancias en las que conviene hacer inferencias causales y asimismo no se pueden diseñar auténticos experimentos, en los cuales la necesidad de una elevada validez interna dificulta la validez externa que se pretende. La validez interna se refiere a la rigurosidad del diseño, mientras que la externa a su poder de generalización. La razón más corriente que no permite usar un diseño experimental es la imposibilidad de asignar al azar los sujetos a los grupos experimentales y de control, o el hecho de no poder disponer de un grupo de control porque no es conveniente o es demasiado costoso. Por suerte, tenemos un buen número de diseños utilizables en estas circunstancias. Estos diseños se denominan cuasiexperimentales porque, aún cuando no son verdaderos experimentos, proporcionan un control razonable sobre la mayor parte de las fuentes de invalidez y son más sólidos que los diseños preexperimentales. Hay muchos diseños cuasiexperimentales, pero sólo presentaremos algunos a modo de ejemplo.

Diseño de pretest-posttest con grupo de control no equivalente

Este diseño es uno de los más utilizados, sobre todo porque a menudo no es posible o conveniente aleatorizar los sujetos. El investigador usa, sin alterarlos, grupos de sujetos ya establecidos; una vez aplicado un pretest, administra un tratamiento a un grupo y posteriormente aplica un posttest. La única diferencia entre este diseño y el de pretest-posttest con grupo de control es la asignación **no aleatoria** de los sujetos a los grupos. Se puede representar como hacemos a continuación:

Grupo	Asignación	Pretest	Tratamiento	Posttest
A	no R	O	X	O
B	no R	O		O

Las amenazas principales a la *validez interna* son que los grupos pueden diferir en lo que se refiere a características relacionadas con la variable dependiente, por ello el investigador orientará tanto como pueda los criterios de formación de los grupos, para descartar hipótesis rivales. Un procedimiento usual consiste en comprobar la equivalencia de los grupos con respecto a la antigüedad como grupo, el tamaño, las calificaciones, las actitudes, el nivel socioeconómico o las puntuaciones del pretest.

Series temporales interrumpidas con un grupo

Este diseño requiere un grupo sobre el cual se hagan muchas observaciones y se obtengan series de puntuaciones, antes y después del tratamiento. Las observaciones previas se pueden considerar como pretests seriados, las posteriores como posttests. Se puede representar así:

Grupo	Observaciones previas	Tratamiento	Observaciones posteriores
A	OOOOOOOOOO	X	OOOOOOOOOO

Se han de reunir diversas condiciones. En primer lugar, las observaciones deben realizarse en períodos de tiempo equivalentes; en segundo lugar, se debe prever una mortalidad experimental baja, y en tercer lugar el tratamiento debe ser claramente identificable y replicable.

Series temporales interrumpidas con grupos de control

La inclusión de un grupo de control mejora el diseño anterior de manera considerable, se elimina la amenaza de la historia –influencia de variables contextuales durante la aplicación del diseño– e incluso la instrumentación. Además, si utilizamos el azar para formar los grupos, la selección no amenaza la validez interna del diseño. Puede representarse del siguiente modo:

Grupo	Observaciones previas	Tratamiento	Observaciones posteriores
A	OOOOOOOOOO	X	OOOOOOOOOO
B	OOOOOOOOOO		OOOOOOOOOO

Como en la mayoría de diseños anteriores, podemos plantear variaciones sobre el esquema básico que hemos presentado. En este caso, por ejemplo, podemos comparar el efecto de diversos tratamientos con un grupo de control y obtener un diseño de series temporales interrumpidas de tratamiento alternativo con grupo de control.

Diseños de sujeto único

Hasta ahora hemos tratado de diseños en los cuales intervienen grupos de sujetos, pero en ciertas ocasiones nos interesa descubrir relaciones de causa-efecto a partir de casos individuales. En este caso, la muestra está formada por un solo sujeto; por eso suele hablarse de **diseño n = 1**. Esta situación es muy frecuente en la investigación sobre sujetos con necesidades educativas especiales. La mayoría de estos diseños implican la valoración continua de determinada conducta durante un período de tiempo, por tanto exigen mediciones múltiples dentro de unas fases previamente determinadas en el estudio. Estas valoraciones continuas permiten extraer conclusiones sobre la eficacia de los programas de intervención aplicados. No es casual que uno de los ejemplos clásicos de estos estudios sea el que explica Dietz (1977) en el que cita un caso único en el que estudiaba el efecto que podía tener un maestro sobre el comportamiento perturbador de un adolescente perturbado mentalmente; las interrupciones molestaban a sus compañeros de clase. El maestro aplicó un programa de refuerzo que consistía en dedicarle un tiempo extra de trato personal si disminuía el número de interrupciones. Antes de la aplicación del tratamiento se hicieron valoraciones continuas del número de interrupciones del alumno. Los resultados de las valoraciones realizadas en esta investigación pueden verse en la figura 5.2. Donde dice "reversal" significa volver a la línea de base (sin tratamiento). SE observa que cuando desaparece el tratamiento aumenta la frecuencia del comportamiento perturbador (hablar alto). Cuando se introduce de nuevo el tratamiento vuelve a disminuir la frecuencia de comportamiento perturbador. En conjunto queda clara la efectividad del tratamiento.

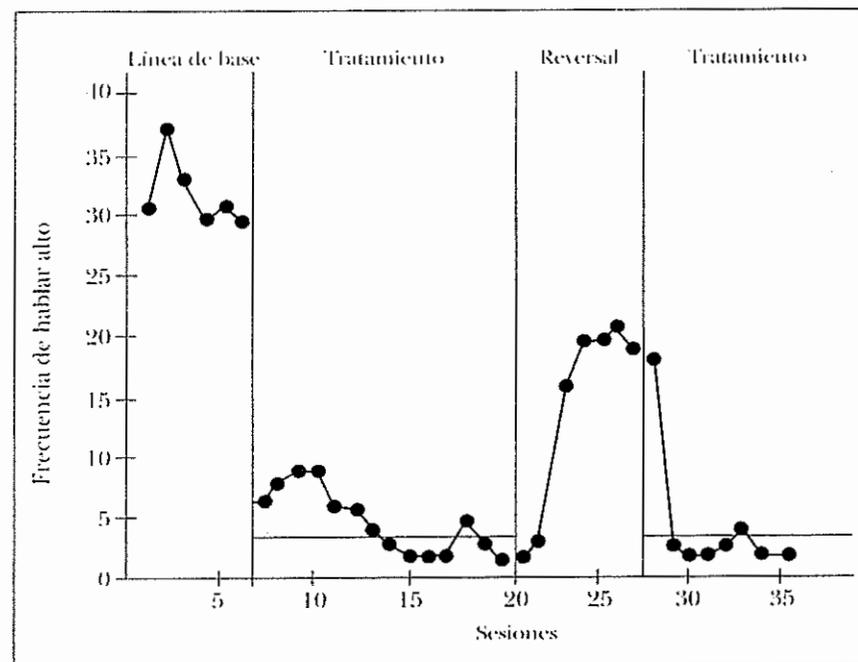


Fig. 5.2.

La fase en la que se valora la conducta del sujeto antes de efectuar ninguna intervención se denomina "fase A", la cual, hecha al principio del diseño se denomina "línea de base" ya que nos indica cual es el estado inicial de la conducta con un cierto grado de estabilidad. También debe observarse que cuando se hacen valoraciones continuas mientras se aplica el tratamiento, la fase se denomina "B". El número y orden de estas fases definen la denominación del diseño. En el ejemplo anterior se trata de un diseño ABAB. Esto significa A (línea de base) B (tratamiento) A (reversal), B (tratamiento). La metodología de sujeto único es una alternativa interesante para determinados tipos de investigación; ofrece ventajas porque respeta la realidad educativa dado que estas observaciones se realizan generalmente en escenarios naturales. En general permite, además de comprobar supuestos teóricos, poner a prueba programas de intervención educativa.

EJERCICIOS

1. ¿Cuál de dichas características no es ninguna de las seis que caracterizan el método experimental?
 - a) Comparación de dos o más grupos.
 - b) Uso de estadísticos descriptivos.
 - c) Control de las variables extrañas.
 - d) Equivalencia de los grupos.
2. Un inconveniente de la aplicación de los métodos experimentales en educación es:
 - a) La dificultad para establecer relaciones causales.
 - b) La poca validez interna.
 - c) El coste excesivo de su aplicación.
 - d) La limitada capacidad de generalizar los resultados.
3. El diseño que ofrece menos validez interna es:
 - a) Sólo posttest con un grupo.
 - b) Pretest-posttest con un grupo.
 - c) ABAB de sujeto único.
 - d) Diseño de series temporales.
4. La realización de un diseño en un proceso de investigación es posterior a:
 - a) La formulación de hipótesis y el análisis de datos.
 - b) La planificación y la recogida de datos.
 - c) La elaboración de conclusiones y la formulación de hipótesis.
 - d) La planificación y la formulación de hipótesis.
5. ¿Cuál es el procedimiento más aconsejable para controlar variables extrañas?
6. ¿Cuántos grupos tenemos que formar para hacer un balanceo completo con las variables "sexo" (hombre, mujer) y "curso de psicopedagogía" (1º y 2º)?
7. ¿Cuántos grupos de sujetos necesitaríamos para realizar un contrabalanceo completo en una situación intrasujeto en la cual interviniese la variable "Método de estudio" (inductivo, deductivo o mixto)?
8. ¿Qué factor puede influir más negativamente en el diseño pretest-posttest con un grupo?

- a) Regresión estadística.
- b) Experimentador.
- c) Testing.
- d) La historia.

CAPÍTULO 6

LA INVESTIGACIÓN EX POST-FACTO

Joan Mateo Andrés

OBJETIVOS

Con el estudio de este capítulo, pretendemos que los alumnos alcancen los objetivos siguientes:

- a) Establecer las diferencias básicas entre la investigación experimental y la ex post-facto.
- b) Conocer las principales metodologías ex post-facto y sus características más importantes.
- c) Saber determinar las posibilidades de aplicación de cada una de las metodologías estudiadas.
- d) Conocer los procesos fundamentales asociados a cada una de las metodologías.
- e) Reconocer el tipo específico de enfoque metodológico ex post-facto que se ha empleado al analizar una investigación.
- f) Saber analizar críticamente los aspectos metodológicos de los trabajos publicados en revistas de investigación educativa.
- g) Conocer los límites y las posibilidades de uso de la investigación ex post-facto en Educación.

INVESTIGACIÓN EX POST-FACTO E INVESTIGACIÓN EXPERIMENTAL

Los fundamentos científicos de la investigación experimental exigen que el investigador controle las condiciones de producción del fenómeno a analizar, como paso a previo al control de las variables que intervienen en el mismo. El control está en la base de la propia lógica de justificación del descubrimiento y en definitiva la validez del experimento depende de nuestra capacidad de control de la varianza.

Sin embargo, la situación más habitual en la investigación en ciencias sociales y humanas, reside en la dificultad de generar y dominar los fenómenos sujetos a estudio, fundamentalmente porque lo más sustantivos acostumbran a producirse al margen de la voluntad del propio investigador. El caso más paradigmático lo constituyen aquellos fenómenos en los que los hechos que los configuran ya se han producido cuando nos aproximamos a su estudio (de donde procede la expresión genérica que califica a este tipo de investigación, *ex post-fact*—después de los hechos—) y difícilmente podremos ejercer su control basándonos en los principios que se aplican en la investigación experimental.

En estos casos, la posición del estudioso y el planteamiento de la investigación se configuran mediante entornos y diseños claramente diferenciados de los experimentales. Su presentación y discusión como una gran alternativa a la investigación más clásica, constituirá el objetivo central de este capítulo.

CLASIFICACIÓN DE LOS ESTUDIOS EX POST-FACTO

A diferencia de la investigación experimental, en la investigación *ex post-facto* no es usual referirse a las diferentes estrategias que se usan en la investigación de esta naturaleza mediante el término “diseño”, en el caso que nos ocupa es más corriente aplicarles el nombre de “estudio”.

Proponemos como clasificación de los diferentes estudios y que consecuentemente nos servirá como modelo organizativo y guía de los apartados en que se subdivide este capítulo, los siguientes:

- A) *Estudios descriptivos.*
- B) *Estudios de desarrollo:*
 - B1. Estudios longitudinales.
 - B2. Estudios transversales.
 - B3. Estudios de cohortes.
 - B4. Estudios de tendencias.
- C) *Estudios comparativo-causales.*
- D) *Estudios correlacionales:*
 - D1. Estudios correlacionales y predictivos.
 - D2. Estudios basados en el análisis factorial.

- D3. Estudios basados en los modelos causales:
 - El análisis de sendas.
 - Los modelos de ecuaciones estructurales.
- D1. Otras aplicaciones multivariadas:
 - El análisis de conglomerados.
 - El escalamiento multidimensional.

LOS ESTUDIOS DESCRIPTIVOS

La investigación en su forma más elemental trata sobre la descripción de fenómenos naturales o debidos a la acción del hombre, así analizar su forma, acción, cambios producidos por el paso del tiempo, similitudes con otros fenómenos, etc., ha constituido a lo largo de la historia el primer paso para conocer de forma sistemática la realidad y debe ser considerada como de vital importancia en los procesos de construcción del conocimiento; está sin duda en la base de muchos logros científicos.

Pensemos, por ejemplo, en el esfuerzo realizado por los astrónomos para desarrollar descripciones respecto de diferentes partes del universo. Este tipo de investigación ha sido capital para muchos otros descubrimientos sobre la estructura de nuestro sistema solar y ha permitido establecer leyes que nos permiten predecir distintos eventos que afectan a los astros.

Aterrizando en nuestro campo del saber, los estudios descriptivos constituyen una opción de investigación cuantitativa que trata de realizar descripciones precisas y muy cuidadosas respecto de fenómenos educativos. Estos estudios son propios de las primeras etapas del desarrollo de una investigación y nos proporcionan hechos, datos y nos preparan el camino para la configuración de nuevas teorías o investigaciones (Fox, 1981).

Como ejemplo podríamos proponer un estudio para detectar el tipo de actividades educativas que desarrollan los adolescentes fuera del periodo escolar, o bien otros sobre la opinión de los profesores respecto de una nueva ley educativa, o bien sobre el estado de la cuestión de los modelos evaluativos que se aplican en los centros educativos, etc. etc.

Centran su actuación en determinar el “qué es” de un fenómeno educativo y no se limitan a una mera recogida de datos sino que intentan responder a cuestiones sobre el estado presente de cualquier situación educativa con implicaciones que

van más allá de los límites establecidos por los propios elementos estudiados.

La recogida de información

Los estudios descriptivos dependen en gran manera de nuestra capacidad técnica para recoger datos sobre los fenómenos que deseamos estudiar. Posiblemente, el gran caballo de batalla del futuro más inmediato será desarrollar tecnologías plausibles que nos ayuden a capturar la información más relevante y significativa de las realidades educativas sobre las que tengamos interés por estudiar y que nos permitan realizar descripciones ricas y substantivas de las mismas.

En la actualidad la investigación educativa se rige por el uso de un conjunto importante de instrumentación, de entre la que destacamos las pruebas estandarizadas de rendimiento, los instrumentos de observación en clase, las escalas de actitud, los cuestionarios, las entrevistas y el análisis de contenido.

Evidentemente no agotamos con este listado el total de las existencias, pero sí que hemos señalado los más ampliamente usados en la investigación descriptiva. Es obvio que un texto de estas características no puede entrar en el estudio detallado de cada una de estas técnicas, que deberían ser objeto de amplio y profundo estudio para cualquier educador.

Fases en un estudio descriptivo

En la planificación de un estudio descriptivo se siguen los pasos habituales de la investigación cuantitativa:

- Identificar y formular el problema a investigar.
- Establecer los objetivos del estudio.
- Seleccionar la muestra apropiada.
- Diseñar o seleccionar los sistemas de recogida de información.
- Recoger y analizar los datos.
- Extraer conclusiones.

Los análisis estadísticos en los estudios descriptivos

Habitualmente para describir una muestra de datos el investigador procede en primer lugar a definir las variables, a medirlas y finalmente a extraer un conjunto de estadísticos que le permiten descri-

bir con cierta precisión la muestra. Estos estadísticos acostumbran a ser las medidas de tendencia central (media, mediana y moda) y las de variabilidad (desviación estándar, varianza y rango). También podemos describir los datos de una muestra apoyándonos en sistemas de presentación gráfica (histogramas, polígonos de frecuencias, etc.) e incluso podemos utilizar formas más sofisticadas y proceder a generar puntuaciones derivadas.

Cualquier curso elemental de estadística nos muestra cómo proceder para aplicar cada una de las técnicas implicadas en los elementos anteriores y consideramos, como en el apartado anterior, que se escapan a las posibilidades de presentación de este capítulo.

LOS ESTUDIOS DE DESARROLLO

Estos estudios tienen como objeto describir la evolución de las variables durante un periodo determinado de tiempo. En general se centran en el análisis de las diferencias asociadas con la edad y, por tanto, pretenden conocer los cambios que se producen en los sujetos con el transcurso del tiempo.

Los trabajos de Jean Piaget sobre el desarrollo infantil se pueden considerar como paradigmáticos en el contexto de este tipo de enfoque. Dentro de los estudios de desarrollo se distinguen a su vez cuatro orientaciones distintas:

- Los estudios longitudinales.
- Los estudios transversales.
- Los estudios de cohortes.
- Los estudios de tendencias.

Los estudios longitudinales

Un estudio longitudinal implica recoger datos de una muestra en diferentes momentos temporales con objeto de analizar cambios o continuidad en las características de los sujetos que componen la muestra. Los estudios longitudinales no son fáciles de llevar a cabo pero resultan esenciales para explorar la naturaleza y los problemas del desarrollo humano

A modo de ejemplo supongamos que estamos interesados en estudiar la evolución de la memoria en niños entre 11 y 16 años. El procedimiento que seguiríamos implicaría en primer lugar

seleccionar una muestra representativa de sujetos de once años a los que se les aplicaría la prueba pertinente para establecer las características básicas de su memoria, prueba que posteriormente iríamos repitiendo de forma periódica, hasta que alcanzaran la edad límite fijada para nuestra investigación. En este momento dispondríamos de la información necesaria para elaborar una ajustada descripción de la evolución de la capacidad estudiada.

Los estudios transversales

Tal como señalábamos los estudios longitudinales exigen seleccionar una muestra y analizar su evolución siguiendo a los individuos a lo largo del tiempo. Evidentemente esta situación genera enormes dificultades en términos de mantenimiento de la muestra (cada año, por una razón u otra, es relativamente fácil que se descuelguen sujetos de la investigación) y en términos de tiempo, ya que cualquier investigación precisa para su ejecución un dilatado plazo temporal.

Para paliar los citados inconvenientes, los investigadores pueden simular condiciones similares a las de los estudios longitudinales evitando sus problemas, mediante el uso de los denominados estudios transversales.

En este tipo de estudio se mantiene el objetivo de estudiar la incidencia del paso del tiempo en el desarrollo de los sujetos pero en lugar de efectuar el seguimiento de una muestra de sujetos de la misma edad a lo largo del periodo de tiempo cuyo efecto interesa investigar, se opta por analizar simultáneamente los componentes de una muestra en la que figuren sujetos de diferentes edades de forma que cubran el arco de tiempo propuesto en el estudio.

Así, en el ejemplo propuesto anteriormente, en lugar de coger una muestra de estudiantes de 11 años para analizar los cambios en la naturaleza de su memoria durante cinco años, escogeríamos cinco grupos más o menos equivalentes en sus características básicas, con excepción en lo que hace referencia a la edad donde deberíamos seleccionarlos de forma que cubrieran las distintas edades comprendidas entre los 11 y los 16 años.

Evidentemente no realizaríamos el seguimiento de los mismos sujetos, con todo lo que comporta de control de las diferencias individuales, pero en cambio no sería necesario esperar cinco

años para finalizar el estudio con todo lo que supone de ahorro en tiempo y esfuerzo.

Los estudios de cohortes

Los estudios de cohortes describen el cambio mediante la selección de muestras distintas para cada momento temporal de recogida de información manteniendo estable la población.

Sin embargo esta definición peca de un exceso de tecnicismo, en realidad el término cohorte se remonta a la época romana y hacía referencia al conjunto de soldados que pertenecían a una misma leva. En nuestro caso una cohorte la forman el conjunto de individuos que comparten algún acontecimiento vital común. De forma genérica este acontecimiento acostumbra ser el año de nacimiento, pero no se excluyen en absoluto hechos de otro tipo. Podríamos considerar como sinónimos del término cohorte otros como promoción, quinta, curso, etc.

A modo de ilustración pongamos como ejemplo el estudio llevado a cabo por Carson, C.C., Huelskamp, R.M. y Woodall, T.D. (1993), que quisieron analizar el grado de abandono de los estudiantes de secundaria; para ello escogieron una muestra nacional (USA) de la cohorte de alumnos que deberían haber finalizado sus estudios en 1982 e hicieron un seguimiento de ellos hasta 1986, tomando información cada dos años.

Si el estudio hubiera sido puntual, hubiera arrojado como resultado del mismo que el número de alumnos que acaban la secundaria, teniendo en cuenta los datos de 1982, es del 82.7% con una tasa de abandono del 17.3%. Sin embargo al recoger información de esta promoción, se observó que había que añadir por un lado un 5.2% más que acabaron en 1984 y otro 2.8% en 1986. Evidentemente esta nueva información de la citada cohorte nos ayuda a comprender con mayor profundidad la naturaleza del abandono escolar.

Los estudios de tendencias

Técnicamente los estudios de tendencias suponen una variante de los de cohortes y describen el cambio como en el caso anterior mediante la extracción de muestras diferentes de sujetos en cada punto temporal de interés para el estudio, pero a diferencia del de cohortes la población no se mantiene ni única ni estable.

Pongamos el siguiente ejemplo para clarificar mejor lo que pretendemos decir: supongamos que deseamos analizar como varían las tendencias en el uso de calculadores gráficos por parte de los profesores de Matemáticas de Bachillerato. Nosotros podríamos definir la población como formada por todos los profesores que constan en el listado de titulados de su colegio profesional. En nuestra investigación procederíamos a extraer una primera muestra el primer año de nuestro estudio y analizaríamos el grado de uso que hacen de este tipo específico de calculadores. Al llegar el segundo año procederíamos nuevamente, pero es algo más que posible que se hubiera modificado la población, dado que cada año se producen altas y bajas; sin embargo procederíamos a seleccionar una nueva muestra y realizaríamos las mismas preguntas y así sucesivamente en años posteriores, para finalmente extraer las correspondientes conclusiones y elaborar el informe de investigación.

Es evidente que se trata de una población que tiene como una característica propia su continua modificación, y también que al variar cada año de muestra introducimos elementos claramente contaminantes en el trabajo, pero aún así se considera que de su análisis surgen conclusiones de estimable valor y que se trata de una situación bastante habitual en la realidad social en la que nos movemos y que consecuentemente deberemos aceptar como razonablemente correcta desde la perspectiva investigadora.

Un estudio típico de tendencias fue el llevado a cabo por la Universidad de Barcelona recogiendo información sobre cómo se había producido la transición desde la finalización de los estudios universitarios al mundo del trabajo expresado en términos de inserción laboral. Se analizaron las promociones de los alumnos que acabaron en los años 1994, 95, 96 y 97, extrayéndose muestras de las cinco respectivas poblaciones y se comprobó entre otras cosas el porcentaje de inserción laboral que se había producido en los alumnos al cabo de dos años de haber acabado la carrera. Este estudio se realizó con casi todas las carreras que se cursan en la mencionada Universidad.

En muchas ocasiones, tanto en los estudios de tendencias como en los de cohortes, se utilizan datos procedentes de otros estudios anteriores que se incorporan para su análisis con otros más actuales recogidos mediante algún tipo de replicación.

Finalmente y a modo de ilustración hemos escogido un estudio propuesto por Bisquerra (1989) que nos parece que puede ayudar a comprender mejor la naturaleza de este tipo de investigación y su aportación a la construcción del conocimiento.

Bisquerra presenta una tabla de doble entrada en la que se recoge la evolución del porcentaje de fumadores (datos obtenidos en los archivos de un centro hospitalario) a lo largo de cuarenta años y teniendo en cuenta un rango de edad comprendido entre los 20 y los 69 años.

Edad	Año				
	1945	1955	1965	1975	1985
20-29	32%	38%	46%	63%	51%
30-39	31%	36%	44%	59%	46%
40-49	28%	32%	41%	56%	48%
50-59	22%	26%	36%	51%	43%
60-69	16%	21%	28%	36%	32%

Observamos que no se trata de un estudio longitudinal, ya que no hablamos de los mismos individuos para cada uno de los puntos de recogida de datos. Tampoco se trata de un estudio transversal, ya que el periodo de observación se prolongó por cuarenta años. Podría tratarse de un estudio de cohortes dado que trabajamos con gente que tiene en común la década de nacimiento, pero al haber juntado muestras procedentes de diferentes poblaciones deberíamos clasificarlo como un estudio de tendencias.

Si analizamos las filas, este estudio nos permite describir el cambio que se ha producido en el hábito de fumar entre fumadores de la misma edad pero considerando el cambio de época.

Si nos fijamos en las columnas podremos observar el cambio existente entre fumadores de diferente edad pero situados en la misma década.

Finalmente si nos situamos en las diagonales, nos encontramos con las cohortes y podemos estudiar como ha ido evolucionando el porcentaje de fumadores de personas pertenecientes a la misma cohorte a medida que iban pasando los años.

Es evidente que este tipo de investigación centra su atención en las diferentes tendencias que aparecen al analizar la información desde las distintas perspectivas que nos ofrece la sabia mezcla de poblaciones y muestras bajo el influjo del paso del tiempo.

Fases en un estudio de desarrollo

- Identificar y formular el problema a investigar.
- Establecer los objetivos del estudio.
- Seleccionar la estrategia metodológica a aplicar (longitudinal, transversal, etc.).
- Seleccionar las poblaciones y muestras correspondientes.
- Diseñar o seleccionar los sistemas de recogida de información.
- Definir el "timing".
- Recoger y analizar los datos.
- Extraer conclusiones.

Los análisis estadísticos

Analíticamente se podrían aplicar diversas técnicas, así podríamos servirnos de curvas de crecimiento, diseños de medidas repetidas, análisis de seguimiento, análisis de series temporales, etc. Una variante especial es aquella que se conoce como diseño de sujeto único $N=1$, en esta variante el estudio se centra en un solo individuo pero lo hace de una forma exclusiva, exhaustiva e intensa.

LOS ESTUDIOS COMPARATIVO-CAUSALES

Los estudios comparativo-causales suponen un tipo de investigación donde el investigador está interesado en identificar relaciones del tipo causa-efecto, más propias de los estudios experimentales, pero que dada la naturaleza del fenómeno resulta imposible por algún motivo, manipular experimentalmente las variables.

Supongamos un caso absolutamente extremo (y absurdo), pero que nos servirá para ilustrar lo que pretendemos explicar. Imaginemos que un equipo de investigadores está interesado en examinar la incidencia que pueda tener sobre la capacidad de resolver problemas en los seres humanos el hecho de que hayan sufrido daños en el cerebro.

Las opciones que "a priori" se le presentan al investigador pueden ser diversas. Podría optar por trabajar con animales inferiores y observar como afecta a su comportamiento cuando se les

daña ciertas partes del cerebro (evidentemente habría que probar la trascendencia del experimento para justificar este tipo de práctica). También deberían resolverse los discutibles paralelismos entre el comportamiento animal y el humano.

Otra posibilidad sería la de experimentar con humanos, pero por razones éticas que no se le escapan a nadie resulta del todo inimaginable y obviamente no vamos a entrar en ello.

Hasta aquí hemos presentado situaciones en las cuales queremos controlar absolutamente la producción del fenómeno, base indispensable para controlar y manipular debidamente las variables, y poder establecer de forma fehaciente la relación causa-efecto.

Sin embargo hemos visto como las variables y el fenómeno, por su naturaleza, imposibilitan este tipo de actuación y es aquí donde puede entrar en juego, para este tipo de situaciones, actuaciones metodológicas mucho más plausibles; en el sentido de que no incurren en problemas éticos y mantienen el rigor y la sistematicidad científica, hablamos de los estudios comparativo-causales.

Bajo la lógica de un estudio comparativo-causal se podría adoptar la estrategia siguiente: en primer lugar, revisar los archivos hospitalarios a la búsqueda de expedientes de personas que hayan sufrido daños en la zonas cerebrales que pretendemos estudiar; en segundo lugar escoger otra muestra de sujetos equivalentes a los anteriores pero que no tengan ningún daño cerebral. Finalmente, contrastar los resultados obtenidos a partir de suministrar las pruebas en las que pongan en juego su capacidad en la resolución de problemas.

Es evidente que cuando se diseña la investigación, los hechos (los daños) ya se han producido y por tanto no hay manipulación de la variable independiente; entramos de pleno en la lógica ex post-facto. La falta de control en la producción del fenómeno impedirá que se pueda establecer "formalmente" la relación causa-efecto, exclusiva de la investigación experimental, pero nadie podrá negar que esta metodología tiene la capacidad de establecer indicios claros de causalidad entre las variables estudiadas.

Un ejemplo, ya real, de este tipo de estudio sería el trabajo de Green y Jaquess (1987) que investigaron el efecto del trabajo a tiempo parcial sobre el rendimiento en estudiantes de secundaria.

Escogieron para ello una muestra de 44 estudiantes de secundaria en la que había estudiantes que trabajan y otros que no. El factor que determinaba la pertenencia a cada uno de los grupos de comparación era evidentemente el hecho de trabajar o no trabajar.

Recogemos en la siguiente tabla los datos que reflejan las puntuaciones de los dos grupos en términos de GPA (puntuación media de las notas):

	<i>Estudiantes empleados Media</i>	<i>Estudiantes no empleados Media</i>	<i>t</i>	<i>p</i>
GPA	2,66	2,78	0,81	0,42

Observamos que en la comparación de medias (prueba t de Student) las diferencias en rendimiento, expresadas en términos de las puntuaciones medias de las notas, no es significativa. A partir de la información analizada no podemos deducir el trabajo a tiempo parcial como un elemento fundamental en la aminoración del rendimiento de los estudiantes de secundaria.

Finalizamos este apartado señalando que en general los metodólogos aceptan que para que pueda existir alguna evidencia clara de causalidad en este tipo de estudio es conveniente que se den las siguientes condiciones:

- Que la secuencia de los hechos se produzca de tal forma que haga posible que A sea la causa de B, e imposibilita que B lo sea de A.
- Cuando se han llevado a términos diversas replicaciones de la investigación ex post-facto, llevadas a cabo por diversos investigadores y surgen resultados consistentes entre todos ellos.

Fases en un estudio comparativo-causal

- Identificar y formular el problema a investigar.
- Definir la población objeto de estudio.
- Seleccionar los grupos de comparación.
- Seleccionar los instrumentos de recogida de información.
- Recoger y analizar los datos.
- Extraer las conclusiones.

Los análisis estadísticos en los estudios comparativo-causales

Fundamentalmente los estudios comparativo-causales implican comparar grupos a través de sus estadísticos más básicos, generalmente las medias. Es por ello que las pruebas estadísticas más utilizadas en este tipo de estudios son:

- Cuando se trata de comparar dos grupos con datos paramétricos se acostumbra a utilizar de forma casi universal la prueba "t" de Student.
- Si seguimos hablando de dos grupos, pero los datos no son paramétricos, se puede optar por la U de Mann-Whitney o la T de Wilcoxon.
- Cuando se pretenden comparar más de dos grupos de forma simultánea, las pruebas más generales son las basadas en el análisis de la varianza. Finalmente existe, para situaciones más complejas, la posibilidad de recurrir al análisis de la varianza multivariable.

Como siempre, cualquier buen manual de estadística nos explicará con detalle las bases y el desarrollo de las mencionadas pruebas.

LOS ESTUDIOS CORRELACIONALES

Estudios correlacionales

Los estudios correlacionales abarcan aquellos estudios en los que estamos interesados en descubrir o aclarar las relaciones existentes entre las variables más significativas de un fenómeno y lo hacen mediante el uso de los coeficientes de correlación.

Un coeficiente de correlación (se puede utilizar indistintamente el término regresión cuando nos referimos a la correlación, ya que matemáticamente la correlación se resuelve mediante un proceso de regresión) es un indicador matemático que aporta información sobre el grado, intensidad y dirección de la relación entre variables. Su cálculo se basa en la suposición de que los datos siguen un modelo lineal teórico, el de la recta de regresión que se expresa en los siguientes términos:

$$Y = a + bX$$

Donde "Y" representa las puntuaciones de la variable dependiente, "a" es el punto de corte de la recta con el eje de ordenadas, "b" representa la inclinación de la recta y "X" los valores de la variable que funcionalmente consideramos como independiente.

El coeficiente de correlación más conocido es el denominado producto-momento o también denominado por el nombre de su

creador: Pearson. Es un coeficiente bivariado, es decir, está diseñado para correlacionar dos variables y fluctúa entre los valores -1 a $+1$. Se representa mediante la letra "r".

Así tenemos:

Correlación	r
Positiva y perfecta	+1
Negativa y perfecta	-1
Nula	0
Positiva e imperfecta	entre 0 y +1
Negativa e imperfecta	entre 0 y -1

Cuando dos variables covarían en la misma dirección existirá una relación que denominaremos positiva y su intensidad quedará reflejada en el coeficiente que adoptará un valor entre 0 y 1, tanto más cerca de 1 cuanto más intensa sea la relación entre las variables.

Para ilustrar lo anterior imaginémosnos la relación entre inteligencia y rendimiento, si midiéramos para un grupo de individuos su inteligencia y su correspondiente rendimiento y con los datos realizáramos el cálculo estadístico de la correlación de Pearson, cabe, por pura lógica, suponer que el cálculo arrojaría un valor que reflejaría una relación positiva de mediana-alta intensidad entre ambas variables.

Un coeficiente cercano a cero nos indicaría la ausencia de relación entre las variables, supongamos como ejemplo que tratáramos de relacionar en un grupo de alumnos su peso y el rendimiento. Lo esperable y lo más probable es que el valor de la relación matemática sería muy bajo, prácticamente cero.

Un alto coeficiente negativo señalaría también la existencia inequívoca de una relación intensa entre las variables pero de direccionalidad opuesta que en términos de puntuaciones se reflejaría en el hecho de que cuando los valores de una de las variables aumenta el de la otra disminuye. Un hecho como este podría ocurrir si correlacionáramos variables como rendimiento y faltas de asistencia a clase.

Actualmente cualquier paquete estadístico nos puede proporcionar este coeficiente, calculado generalmente con cuatro decimales, indicando su grado de significación y acompañado de un diagrama de dispersión, que no es sino un gráfico con todos los pares de observaciones de cada sujeto introducidos en el cálculo

del coeficiente representados por puntos situados en un eje de coordenadas. Este gráfico facilita la inspección de los datos y el posible descubrimiento de relaciones de las variables que nos son estrictamente rectilíneas.

Otros coeficientes bivariantes de correlación

Además del coeficiente de Pearson, que es el más conocido, existen otros índices para el establecimiento de relaciones bivariantes. La decisión de utilizar uno u otro dependerá fundamentalmente de la naturaleza de las variables y en algún caso de la falta de linealidad de la relación. A modo de resumen ofrecemos el cuadro siguiente:

Tipo de Correlación	Símbolo	Variable 1 (naturaleza)	Variable 2 (naturaleza)	Anotaciones
Producto-momento (Pearson)	r	continua	continua	Es la técnica más estable
Kendall	tau	ordinal	ordinal	Preferible Para $n < 10$
Biserial	r bis.	dicotomizada	continua	Usada en el análisis de ítems.
Biserial-puntual	r pbis.	dicotómica	continua	Produce correlaciones inferiores que r bis.
Tetracórica	rt	dicotomizada	dicotomizada	Usada cuando es posible dicotomizar las variables.
Coefficiente fi	fi	dicotómica	dicotómica	Usada en el cálculo de correlación entre ítems.
Coefficiente eta	eta	continua	continua	Usada para reflejar relaciones no lineales.

Los coeficientes de correlación múltiple

Hasta ahora nos hemos limitado a presentar los diferentes índices pensados para explorar relaciones bivariadas. Sin embargo el modelo puede expandirse en el sentido que podemos considerar la posibilidad de que existan fenómenos que su estudio implique simultáneamente (más de dos variables) y precisemos coeficientes que sean capaces de capturar este nuevo tipo de relación multivariable. Hablamos de la correlación múltiple.

El coeficiente de correlación múltiple es un índice que expresa la relación existente entre un conjunto de variables que funcionalmente consideramos como independientes respecto de otra que le otorgamos la categoría de dependiente. Este coeficiente se expresa mediante la letra "R".

A modo de ejemplo supongamos que mantenemos la hipótesis de que el rendimiento escolar está condicionado por la capacidad intelectual del alumno pero también por sus hábitos de estudio. Nos encontramos con un estudio en el que intervienen tres variables, dos que podríamos denominar como independientes –capacidad intelectual y hábitos de estudio– y otra que funcionalmente actúa como dependiente (ya que se supone desde consideraciones teóricas que su resultado depende de las otras dos), que es el rendimiento.

Desde el punto de vista del modelo matemático expresaríamos la relación anterior mediante la siguiente ecuación:

$$Y = b_1X_1 + b_2X_2 + a$$

Donde Y representa los valores de la variable dependiente (rendimiento), X_1 una de las variables independientes (inteligencia), X_2 la otra variable independiente (hábitos de estudio), a es una constante que en el supuesto que trabajáramos con variables estandarizadas equivaldría a cero y los coeficientes b son valores que ponderan la aportación de cada una de las variables al modelo.

El proceso de cálculo busca en primer lugar la matriz de correlaciones entre todas las variables del estudio y posteriormente y a partir de estos valores se genera un sistema de ecuaciones cuya resolución nos facilita los valores de las constantes del modelo. Sin entrar en más disquisiciones matemáticas, que no serían propias del enfoque de este manual, volvemos a recordar

que los programas estadísticos informatizados resuelven fácilmente los cálculos correlacionales. Básicamente introduciendo la matriz de datos de las puntuaciones de los sujetos en todas las variables el programa genera toda la información precisa para interpretar el estudio correlacional.

Conviene recordar que el coeficiente de correlación múltiple va de 0 a +1 y no puede ofrecer nunca valores negativos. Fuera de este detalle, la interpretación sigue las mismas pautas que las que señalábamos en la correlación de Pearson.

Hasta aquí hemos descrito el modelo más común de correlación múltiple; usualmente se les conoce como el modelo ordinario de regresión basado en los cuadrados mínimos. Existen, sin embargo, muchas otras variantes de correlaciones de entre las que destacamos las siguientes:

La correlación parcial es un submodelo dentro de la regresión múltiple, que permite establecer el grado de relación entre dos variables después de controlar la influencia de una tercera.

La correlación canónica constituye la máxima expansión posible del modelo de regresión múltiple; en él se plantea la situación de correlacionar un conjunto de variables independientes frente a otro de variables dependientes. (a diferencia de la múltiple en que la variable dependiente es únicamente una).

El análisis discriminante es una técnica que se emplea para distinguir diferentes grupos que habían sido definidos previamente sobre la base de las puntuaciones obtenidas en una o más variables. Técnicamente es una correlación múltiple en la que la variable dependiente es categórica frente a las independientes que son continuas.

Correlación curvilínea es un modelo de regresión que permite resolver el caso en que la relación entre las variables no es lineal.

Cabe señalar que para entrar en estas disquisiciones se precisa no un manual de estadística de tipo general sino alguno mucho más especializado como podría ser el dedicado al análisis multivariable de Bisquerra (1989).

Interpretación del coeficiente de correlación

Para valorar la intensidad de los índices de correlación, lo habitual es hacerlo calculando su significación estadística. Un

estadístico, como por ejemplo la propia r de Pearson, es significativo cuando se puede afirmar, aceptando una determinada probabilidad de error, que es estadísticamente distinto de cero. Su determinación es muy sencilla y las rutinas que implica su cálculo se pueden encontrar en la mayoría de los manuales de estadística. En la actualidad los programas estadísticos incorporan los mencionados cálculos y nos facilitan el grado de significación expresándolo en términos de probabilidad.

Sin embargo y a efectos puramente prácticos, puede ser útil tener presente un cuadro orientador como el siguiente, basado en lo que es habitual en los estudios de correlaciones en Ciencias Sociales.

<i>Coefficiente</i>	<i>Interpretación</i>
De 0 a 0.20	correlación prácticamente nula
De 0.21 a 0.40	correlación baja.
De 0.41 a 0.70	correlación moderada.
De 0.71 a 0.90	correlación alta.
De 0.91 a 1	correlación muy alta.

Finalmente añadiremos, respecto de la interpretación de los coeficientes de correlación, una última manera muy intuitiva de hacerlo y es mediante el denominado coeficiente de determinación.

El coeficiente de determinación es el cuadrado del coeficiente de correlación multiplicado por 100. El resultado que así se obtiene se interpreta como el porcentaje de varianza (variabilidad) que la variable independiente explica de la dependiente. Así en nuestro caso en que hablábamos de inteligencia y rendimiento y también considerando que funcionalmente podemos hablar de la inteligencia como independiente y el rendimiento como dependiente, si correlacionaran 0.70, podríamos considerar que la inteligencia explica el 49% del rendimiento (0.7 al cuadrado y multiplicado por 100).

El diseño de investigación en los estudios correlacionales

El diseño básico en la investigación correlacional es muy simple, implica únicamente recoger datos de dos o más variables para un conjunto de sujetos y proceder estadísticamente a calcular la correlación. En realidad el gran éxito y aceptación que tiene entre los investigadores esta opción metodológica se debe a la

simplicidad de su diseño como estrategia investigadora, a la sencillez de los cálculos de los distintos coeficientes de correlación existentes y a la simplicidad interpretativa de los resultados obtenidos.

Sin embargo y a pesar de su simplicidad, hemos de recordar que la calidad de una investigación no depende de la complejidad de su aparato metodológico o de la sofisticación de sus técnicas de análisis sino de la profundidad y racionalidad de los planteamientos teóricos que guían el diseño de investigación.

Como colofón a esta última observación y para evitar ciertas alegrías en el uso de los planteamientos correlacionales (un investigador poco cuidadoso puede usar de forma irreflexiva la correlación seleccionando de forma poco ortodoxa las variables, animado en ocasiones por la propia simplicidad del diseño) consideramos como una recomendación fundamental la de tener siempre presente que para incrementar la probabilidad de alcanzar resultados significativos en una investigación correlacional es imprescindible, para orientarla debidamente y seleccionar correctamente las variables más substantivas, reflexionar en profundidad desde la propia teoría y apoyarse en estudios anteriores.

Correlación y causalidad

La aproximación correlacional, al igual que los otros estudios ex post-facto, no permite el establecimiento formal de la relación de causalidad. El que dos variables correlacionen (por ej. Autoconcepto y Rendimiento) quiere decir simplemente que covarían pero no podemos inferir alegremente que una es la causa de la otra. Caben muchas posibilidades de error en la interpretación que el diseño por sí mismo no puede solucionar. Por ej. que la dirección sea la contraria de la que hipotetizamos (que la causa del autoconcepto sea el rendimiento), o que interactúen y no se dé una sola direccionalidad, o que exista una tercera variable que sea realmente la causa y esté impregnando la acción de la que estamos tentados de considerar como causante, etc.

Es por todo ello que deberemos ser muy cautos en las explicaciones derivadas del uso de la correlación. Nuevamente hemos de referirnos a la importancia de la teoría, estudios anteriores y la reflexión crítica para guiar la estrategia correlacional.

Fases en un estudio correlacional

- Identificar y formular el problema a investigar.
- Establecer los objetivos del estudio.
- Seleccionar las variables a correlacionar.
- Diseñar o seleccionar los sistemas de recogida de información.
- Recoger y analizar los datos.
- Interpretar los resultados a la luz de la teoría.
- Extraer conclusiones.

Los estudios predictivos

Los estudios predictivos constituyen una variante dentro de los correlacionales muy apreciada por los investigadores en educación. Fundamentalmente la mayoría de los estudios predictivos suministran tres tipos de información:

- Determinar hasta que punto un patrón de conducta puede ser predicho.
- Suministrar datos para el desarrollo teórico acerca de los determinantes del citado patrón de conducta.
- Presentar evidencias de validez predictiva de una prueba mediante la correlación de las puntuaciones de los sujetos en la misma y el patrón de conducta utilizándolo como variable criterio.

Este tipo de estudios ha proporcionado interesantes trabajos en lo que hace referencia a predecir el éxito de los estudiantes en algún tipo de estudio basándonos en el conocimiento previo que tenemos de un conjunto de sujetos en variables que presumiblemente son las causas de su buen funcionamiento en los mencionados estudios.

Generalmente el investigador se plantea este tipo de estudio cuando después de haber estabilizado un coeficiente de correlación y el conjunto de variables del modelo lineal que lo conforman y valorando como altamente significativo el grado del coeficiente, decide plantearse la situación no tanto desde la perspectiva de explorar y estudiar el tipo y la intensidad de las relaciones entre las variables sino con ánimo de poder predecir conductas

posteriores de los sujetos a partir de la información que sobre ellos podemos recoger previamente.

A modo de ilustración y siguiendo con el ejemplo planteado en la correlación múltiple, respecto a si existe una estrecha relación entre la inteligencia (X_1) y los hábitos de estudio (X_2) por una parte y el rendimiento (Y) por otra, supongamos que después de los correspondientes cálculos, obtenemos la siguiente matriz de correlaciones:

	Y	X_1	X_2
Y	1.00	0.849	0.781
X_1		1.00	0.697
X_2			1.00

Que se interpreta como que entre el rendimiento y la inteligencia existe una correlación de 0.842, entre el rendimiento y los hábitos de estudio el coeficiente es de 0.781 y entre las dos variables independientes es decir inteligencia y hábitos de estudio la relación es de 0.697.

Esta tabla reproduce una situación muy habitual en los estudios realizados en fenómenos propios de las Ciencias Sociales, en los que difícilmente las variables son absolutamente independientes entre sí. En nuestro caso vemos que la correlación por separado entre inteligencia y rendimiento es alta, como también lo es entre hábitos de estudio y rendimiento. Si buscásemos la correlación múltiple, tratando de unir la fuerza de las dos variables explicativas frente a la que pretendemos explicar, cabría pensar que la nueva correlación sería mucho más alta. Sin embargo y dado que las dos variables explicativas (inteligencia y hábitos de estudio) están muy correlacionadas entre sí lo que ocurrirá es que apenas se incrementará el valor de la correlación simple.

Siguiendo con nuestro caso y aplicando los datos proporcionados por el ordenador, la denominada ecuación de regresión (modelo lineal una vez operativizado) sería la siguiente:

$$Y = 0.513X_1 + 0.485X_2 - 0.426$$

Que nos ofrecería una correlación múltiple de $R=0.9$

Con toda esta información lo que básicamente conocemos es:

- Que la correlación múltiple entre inteligencia y hábitos de estudio por un lado y rendimiento por otro es muy alta. (0.9).
- Que la decisión de introducir dos variables independientes en el modelo no ha incrementado sustancialmente la correlación, ya que la inteligencia por sí sola ya correlacionaba 0.849 con el rendimiento.
- Que posiblemente la causa del poco incremento es que inteligencia y hábitos de estudio, tal como se ha planteado en nuestro estudio, no son dos variables autónomas entre sí (su correlación es de 0.697).
- Que, sin embargo, dado el alto coeficiente de correlación podemos atrevernos (con todas las prevenciones correspondientes) a pasar a un estudio predictivo con la información que se posee.

Así, una vez establecida la ecuación de regresión, podemos usarla en planteamientos del tipo siguiente:

Supongamos que al curso siguiente de finalizar el estudio, tenemos un nuevo conjunto de alumnos, de los que conocemos desde el principio su puntuación en inteligencia y en hábitos de estudio y quisiéramos anticipar cual sería su probable nota de rendimiento al acabar el mismo, y de esta forma tratar de enfocar mejor nuestra tutoría y procesos de orientación educativa.

Podríamos, a partir de la información del estudio anterior, situar la información que ahora poseemos de los nuevos alumnos en la ecuación de regresión averiguada y estimar la posible nota de rendimiento.

Así, si un alumno posee una puntuación de 7 puntos de inteligencia y 6 en hábitos de estudio, la estimación de su rendimiento antes de que el curso se inicie sería:

$$Y = 0.513 (7) + 0.485 (6) - 0.426 = 5.975$$

Casi un 6 sería lo previsible para este tipo de alumno en función de la información analizada. Es evidente que se trata de una pura estimación, pero en educación tratar de predecir y anticiparnos a los hechos es el medio más eficaz para prevenirlos y evi-

tar posibles fracasos. Somos conscientes que en el fondo se trata de una conjetura respecto de lo que puede ocurrir, pero se trata de una conjetura basada en hechos y en estudios sistemáticos y bien merecen nuestra más alta consideración y nos ayudan, siempre bajo la reflexión crítica y la guía del análisis teórico más riguroso, a construir conocimiento científico en nuestras áreas de actuación.

LOS ESTUDIOS BASADOS EN EL ANÁLISIS FACTORIAL

Con el término análisis factorial hacemos referencia a un conjunto de técnicas estadísticas que tienen como objeto proveer las bases empíricas para reducir un gran grupo de variables a un número menor de variables hipotéticas denominadas factores.

Técnicamente el análisis factorial combina aquellas variables que están alta o al menos moderadamente correlacionadas entre sí, formando a partir de ellas un factor, que es la expresión matemática del elemento común que tienen las variables combinadas.

Su propósito general es resumir la información contenida en un número de variables originales en un conjunto menor de nuevas dimensiones que llamamos factores, tratando de perder la mínima información posible.

Fases en un análisis factorial

El desarrollo de un análisis factorial sigue las cuatro fases siguientes:

- Cálculo de la matriz de correlaciones entre todas las variables a partir de los datos originales.
- Extracción de los factores
- Rotación de los factores con la intención de facilitar la información.
- Cálculo de las puntuaciones factoriales.

La matriz de correlaciones

El primer paso en un análisis factorial es el de correlacionar entre sí todas las variables que participan en el estudio y cuyo gran número deseamos reducir a unos pocos factores sin perder demasiada información.

En términos técnicos este paso supone pasar de la matriz general de datos a la matriz de correlaciones.

La extracción de factores

El proceso de cálculo nos lleva a una segunda fase en la que empieza propiamente a actuar el análisis factorial y es la de intentar agrupar las variables en factores; a este momento se le denomina extracción de factores; matemáticamente equivale a un proceso de descomposición de la matriz de correlaciones en dos matrices: la factorial y su inversa.

En el contexto del análisis factorial se acostumbra a aplicar dos métodos para resolver la mencionada descomposición: el del análisis factorial común y el de los componentes principales.

La diferencia fundamental entre un método u otro radica en el valor que colocan en la diagonal mayor de la matriz de correlaciones, si conservan los unos estamos hablando del análisis de componentes principales y si son sustituidos por las "comunalidades" (varianza común entre todas las variables del estudio), nos situamos en el análisis factorial común.

La elección de uno u otro método acostumbra a ser un tema menor, pues los resultados no difieren sustancialmente, pero si básicamente lo que se desea es extraer el mínimo número de factores que se necesitan pero explicando la mayor varianza posible de la contenida en las variables originales, la opción a seguir es la del análisis de componentes principales. Si el objetivo primordial es identificar la estructura latente implícita en las variables originales, la opción debería ser el análisis factorial común.

La rotación de factores

Una vez extraídos los factores lo que tenemos en esencia es una nueva matriz de correlaciones (que ahora en este nuevo contexto se denominan "saturaciones") entre cada una de las variables del estudio y los factores extraídos.

Sin embargo esta matriz representa simplemente una de las múltiples soluciones matemáticas que puede adoptar el análisis factorial y no siempre la primera que hallamos, o mejor dicho, que nos ofrece el ordenador, tiene sentido desde la perspectiva de la naturaleza de la investigación que estamos realizando.

Para que la matriz factorial nos ofrezca básicamente lo que estamos buscando es necesario que cada variable esté significativamente saturada por un solo factor. Para conseguir este propósito se procede a lo que se conoce como "rotación de los factores", que gráficamente consiste en hacerlos girar sobre sus ejes de coordenadas de forma que se sitúe en el espacio de manera que cumpla matemáticamente el principio mencionado al principio de este párrafo.

Los métodos de rotación utilizados en el análisis factorial se subdividen en dos básicos: los denominados métodos de rotación ortogonal y los de rotación oblicua. En el primer método los factores finalmente extraídos cumplen el principio que son independientes entre sí (en términos estadísticos se traduce en que entre ellos la correlación es cero), mientras que en la rotación oblicua los factores están correlacionados.

Las puntuaciones factoriales

Una vez reducido el conjunto de variables originales a un número sustancialmente menor de factores, puede ser que surja el interés por conocer cuáles serían las puntuaciones de los sujetos en estos factores.

El cálculo de las puntuaciones factoriales se realiza a partir de la matriz factorial y para efectuarlos se aplican los mismos principios que habíamos visto en la regresión múltiple, cuando estimábamos la puntuación de un sujeto a partir de la información que de él poseíamos sobre sus puntuaciones en las variables independientes.

Generalmente todos los programas de ordenador nos ofrecen directamente las puntuaciones factoriales para cada uno de los sujetos de la muestra y no es necesario efectuar el cálculo.

A modo de ejemplo

Nos serviremos del siguiente ejemplo para tratar de ilustrar lo hasta ahora explicado:

Supongamos que hemos administrado seis pruebas de rendimiento a un grupo de alumnos de doce años y queremos ver si con un número menor de pruebas se podría alcanzar un conocimiento global parecido de estos alumnos. Dicho en otros términos: queremos estudiar si es posible reducir el número de variables perdiendo la menor información posible.

Las pruebas aplicadas fueron las siguientes: Vocabulario (V), lectura (L), sinónimos (S), números (N), aritmética (A) y álgebra (AL).

Con las puntuaciones de la muestra de alumnos en todas las pruebas se procedió a correlacionar las pruebas obteniéndose la siguiente matriz:

	V	L	S	N	A	AL
V	1	0.72	0.63	0.09	0.09	0.00
L		1	0.57	0.15	0.16	0.09
S			1	0.14	0.15	0.09
N				1	0.57	0.63
A					1	0.72
AL						1

La descomposición de la anterior matriz nos permitió extraer los factores expresados en la siguiente matriz factorial, que como se puede comprobar, al no estar rotada, no cumple con el principio de que cada variable ha de procurarse que sea saturada significativamente por un solo factor:

Variables	Matriz factorial no rotada	
	Factores	
	I	II
V	0.60	-0.58
L	0.63	-0.49
S	0.56	-0.43
N	0.56	0.43
A	0.63	0.49
AL	0.60	0.58

Finalmente se procedió a una rotación ortogonal para tratar de dotar de sentido a la matriz:

Variables	Matriz factorial rotada	
	Factores	
	I	II
V	0.83	0.01
L	0.79	0.10
S	0.70	0.10
N	0.10	0.70
A	0.10	0.79
AL	0.01	0.83

Del análisis de esta última matriz se desprende cómo las tres variables primeras que hacían referencia a pruebas del entorno lingüístico son las que saturan al factor I, mientras que las otras tres del ámbito matemático, saturan al factor II. Es evidente que el conjunto de seis pruebas se podría reducir a dos, quizás algo más amplias, que contemplaran: la primera, cuestiones lingüísticas y la otra, matemáticas.

No siempre el espectro que aparece es tan evidente y sencillo como éste. El ejemplo que hemos presentado es puramente artificial y pensado para ilustrar con sencillez la explicación algo compleja del análisis factorial. Pero aún así la técnica aquí presentada acostumbra a comportarse de forma tremendamente eficaz y sus resultados son habitualmente muy sugerentes y creativos. Se considera una técnica muy valiosa en los procesos de generación de teoría y es por ello que en las nuevas aportaciones han aparecido dos enfoques: uno más orientado a la exploración de las relaciones entre variables y otro dirigido a la confirmación de sus estructuras latentes, aunque sin embargo la discusión de estos temas ya es objeto de manuales más especializados sobre el análisis factorial.

LOS ESTUDIOS BASADOS EN EL ANÁLISIS CAUSAL.

Los modelos causales conforman un conjunto de estrategias que tienen por objeto principal validar modelos causales elaborados por el investigador y fundamentados en una teoría.

Los orígenes se encuentran en el análisis de sendas ("path analysis") del genetista Wright (1934) y en las ecuaciones estruc-

turales de los econométricos de los años cuarenta, pero la formalización matemática del análisis (Duncan, 1975) y especialmente su conversión en programas para ordenador (Joreskog y Sörbom, 1985) son relativamente recientes.

El análisis de sendas

El análisis de sendas es un método para probar la validez de una teoría respecto a las relaciones causales entre tres o más variables que han sido previamente analizadas mediante un estudio correlacional.

Fases en un análisis de sendas

El análisis de sendas contempla tres fases:

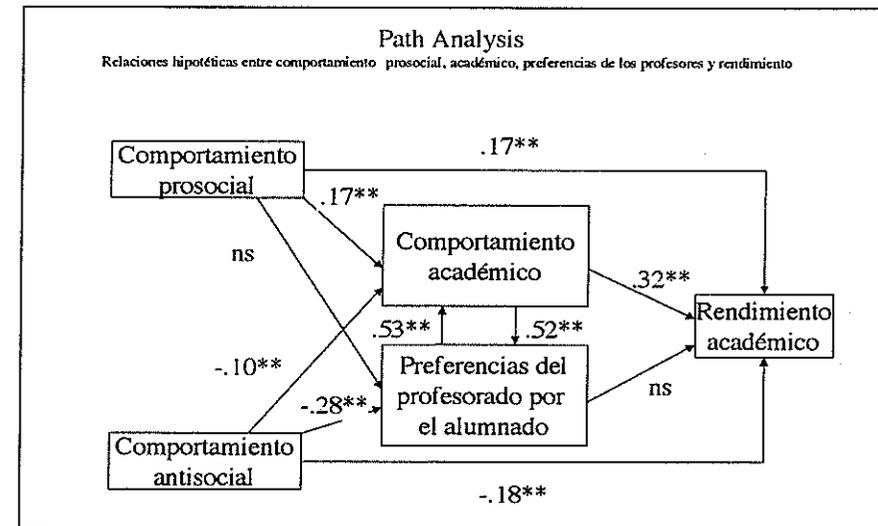
1. Consiste en formular hipótesis respecto de las conexiones causales existentes entre las variables que intervienen en el estudio.
2. Se seleccionan o desarrollan medidas de las variables (también se les denomina en este contexto: constructos teóricos) que previamente habíamos especificado en la fase anterior. Este paso es fundamental, ya que si la calidad de la información recogida sobre una variable es baja, resultará imposible alcanzar resultados que impliquen tan alto compromiso con los procesos de generación de teoría.
3. Calcular los correspondientes estadísticos que nos confirmen o nos refuten la fortaleza de las relaciones entre todos los pares de variables en el marco de las relaciones causales hipotetizadas en el modelo teórico.

A modo de ejemplo

Para ilustrar brevemente lo anterior proponemos seguir el trabajo de Wentzel (1993), en el que trató de determinar los efectos del comportamiento social y antisocial de los alumnos en su rendimiento. Wentzel deseaba determinar si los efectos del comportamiento social afectaban directamente a los resultados o bien previamente afectaban las preferencias de los profesores por sus alumnos y/o a la propia orientación académica de los alumnos.

Para resolver su investigación en primer lugar resolvió el tema de las medidas de las variables y posteriormente para una amplia muestra de alumnos recogió la información pertinente; posteriormente intercorrelacionó todas las variables.

Es a partir de la matriz de correlaciones y de los supuestos teóricos del modelo causal que se pueden calcular los coeficientes definitivos (se denominan efectos) de las sendas entre las variables. Estos coeficientes reducen la correlación inicial al valor que realmente es atribuible a la relación entre variables cuando se tiene en cuenta la totalidad del modelo. En el caso del estudio que hemos mencionado, el modelo quedó así:



En el modelo que aquí presentamos se observa que todas las flechas tienen básicamente una sola dirección (así, la conducta social incide en la académica, representado por la flecha pero no se teoriza respecto del camino contrario). Este modelo así diseñado recibe el nombre de recursivo. Si el modelo considerara en sus hipótesis relaciones recíprocas y no unidireccionales (como es el caso), le denominaríamos no-recursivo.

Otras dos expresiones que conviene recordar es la de variables exógenas y endógenas. Así, la conducta social y la antisocial se consideran en este estudio como exógenas por que no estamos

hipotetizando que reciban la influencia de ninguna otra variable. Por el contrario, el resto de las variables son endógenas porque hipotetizamos respecto de ellas que, al menos, reciben una influencia de otra variable.

Finalmente en lo que concierne a la interpretación de este modelo se puede observar como la incidencia directa de la conducta social o antisocial del alumno no influye directamente en el rendimiento (la correlación inicial se ha diluido en unos exiguos efectos de sólo .17 y -.18) sino que lo hace a partir de la influencia indirecta que tiene tanto sobre la percepción de los profesores como de la propia orientación de la conducta académica de los alumnos, donde sí se mantienen importantes coeficientes causales.

Los modelos de ecuaciones estructurales

Existe un modelo más potente y sofisticado para analizar las relaciones causales entre variables, se trata del modelo denominado de ecuaciones estructurales, más conocido por el nombre de LISREL (acrónimo de "Lineal Structural Relationships), que es la denominación del programa para ordenador generado para solucionar los cálculos que implica.

Este tipo de análisis resuelve mejor que el análisis de sendas el proceso de validación del modelo, aunque resulta más engorroso su cálculo y más compleja su interpretación.

Fundamentalmente la diferencia radica en que mientras en el análisis de sendas sólo podemos probar la significación de los efectos entre cada par de variables y podemos teorizar respecto de la conveniencia de mantenerlos o eliminarlos del modelo, los modelos de ecuaciones estructurales nos ofrecen coeficientes de ajustes generales para todo el modelo en su conjunto, al margen naturalmente de ofrecer también los efectos entre variables como hacía el de sendas.

Dado por un lado la mejora que supone este modelo respecto del anterior y la cada vez mayor accesibilidad al uso de los programas de ordenador, parece que los modelos estructurales acabarán convirtiéndose en la única referencia de este tipo de análisis.

Fases en los modelos de ecuaciones estructurales

De forma muy breve, ya que entrar en profundidad exigiría planteamientos mucho más sofisticados, vamos a presentar los pasos que sigue el investigador en el análisis:

- Especificación del modelo.
- Identificación del modelo.
- Estimación de parámetros.
- Evaluación del modelo.

Especificación del modelo

El primer momento del análisis causal es el de elaborar un modelo teórico que posteriormente tendremos que validar. Tal como hemos visto en el ejemplo anterior (el de la conducta social y antisocial), el modelo se representa mediante un diagrama de flujo.

La identificación del modelo

El modelo se ha de formular en forma de ecuaciones estructurales y hemos de asegurarnos que tiene una única solución. Entrar en las matemáticas que implica este proceso de identificación se escapa absolutamente de las posibilidades de este manual; el lector puede, sin embargo, consultar los trabajos de Duncan (1975), si desea profundizar en el tema.

En general todos los modelos recursivos que incorporan las variables más relevantes, tienen la identificación prácticamente asegurada.

La estimación de parámetros

Tras la identificación del modelo y con todos los datos sobre las variables ya recogidos, se procede a introducirlas en algún programa de ordenador que nos resuelva las ecuaciones planteadas y nos estime los parámetros más básicos del modelo para poder proceder a su validación. El programa de mayor difusión es el anteriormente mencionado, el LISREL (Jöreskog y Sörbom, 1985) que está actualmente incorporado en los paquetes de la serie SPSS-X donde figura en su anexo).

La evaluación del modelo

Después de estimar los parámetros se procede a la validación del modelo que consiste en comprobar si el modelo que habíamos hipotetizado se ajusta a los datos obtenidos empíricamente. La

comprobación del ajuste se realiza mediante pruebas similares a las de X cuadrado.

En el supuesto de que no ajustara, se puede proceder a una modificación del modelo inicial, si entendemos, tras su análisis, que aún es posible con algún reajuste de las flechas o de las variables que intervienen el salvarlo, de lo contrario habría que abandonarlo e iniciar de nuevo la investigación. Una vez modificado se puede proceder a un reanálisis de los datos y probar nuevamente su validez.

Se entiende que cuando conseguimos el ajuste, el modelo teórico ha quedado definitivamente validado.

OTRAS APLICACIONES MULTIVARIABLES

Resultaría muy prolijo el tratar de presentar todas las posibles aplicaciones multivariantes y, volvemos a insistir, no es el objetivo de un manual generalista como el que aquí desarrollamos. Sin embargo, sí que de forma muy sucinta nombraremos, para acabar nuestro capítulo, dos de las más conocidas en la investigación en ciencias sociales.

El análisis de conglomerados

Este análisis hace referencia a un conjunto de técnicas, fundamentalmente algoritmos, que sirven para agrupar individuos o variables en clases o conglomerados.

El análisis de conglomerados se conoce también como análisis Q, análisis tipológico, análisis de clasificación y taxonomía numérica. Esta gran variedad de nombres se debe en gran parte al uso diverso que el análisis de conglomerados ha tenido en disciplinas tan distintas como puedan ser la Biología, la Educación, la Empresa, la Psicología, etc. Sokal y Sneath (1963), dos de los autores que más han influido en el desarrollo de este modelo, consideran que la clasificación es uno de los procesos más fundamentales de la ciencia, ya que los fenómenos han de estar ordenados para que los podamos comprender. La importancia del análisis de conglomerados consiste en su capacidad para clasificar los individuos o las variables en categorías.

Para aplicar el análisis de conglomerados se procede en cuatro fases:

- La selección de las variables; este es un paso fundamental, es muy importante detectar las que son realmente relevantes para el proceso de identificación de los grupos.
- Determinación de la medida de proximidad más pertinente. Para la formación de conglomerados (conjuntos de variables que pertenecen a una misma clase), es necesario encontrar un criterio para fijar las distancias entre individuos y poder, con estas mediciones, generar la correspondiente matriz de datos.
- Agrupación de los individuos. Para la agrupación de los individuos el análisis de conglomerados procede con dos metodologías distintas: los métodos jerárquicos y los no jerárquicos. Los más usados son los jerárquicos.
- Presentación de los resultados. Los resultados, una vez analizada la matriz de proximidades mediante la metodología que hayamos decidido aplicar, se presentan de manera analítica en forma de tablas numéricas, pero es muy habitual escoger también algún tipo de presentación gráfica de cómo se han clasificado finalmente los individuos. El gráfico más habitualmente usado es el dendograma, que no es más que un árbol lógico de fácil lectura.

Si se desea ver algún ejemplo de este tipo de investigación puede consultarse entre otros el trabajo de Sans (1989) donde en un estudio sobre los resultados de la selectividad en Cataluña, identificó cuatro tipologías de centros perfectamente diferenciados.

El escalamiento multidimensional

El escalamiento multidimensional hace referencia a un conjunto de técnicas que utilizan como datos medidas que reflejan las proximidades existentes entre cualquier clase de objetos (estas proximidades pueden ser también preferencias expresadas por un grupo de sujetos respecto de los objetos mencionados).

La misión del escalamiento consiste en identificar si existen dimensiones clave subyacentes a las proximidades calculadas y a partir de esta información probar de situar los objetos dentro de un espacio multidimensional.

Es evidente que la definición queda muy abstracta y por ello vamos a servirnos del ejemplo-recurso, ya clásico, que acostum-

bran a incluir (con alguna que otra variante como haremos nosotros) casi todos los manuales que hablan sobre esta técnica (ver Kruskal y Wish, 1984: 8).

Supongamos que tenemos un mapa de España en el que están situadas las ciudades más importantes. Si alguien nos pidiera que construyésemos una tabla con las distancias que hay entre todas ellas, tan sólo necesitaríamos una regla y conocer la escala del mapa, y fácilmente podríamos contestar a la cuestión.

Consideremos ahora el mismo problema pero a la inversa, es decir, tenemos la tabla con las distancias (proximidades) entre las ciudades y nos piden que reproduzcamos el mapa; pues, bien en esencia el escalamiento multidimensional es una técnica que resuelve el problema tal como lo tenemos planteado ahora.

Una proximidad es un número que nos indica en que medida dos objetos son similares o diferentes. El análisis nos ofrece una representación espacial que consiste en una configuración geométrica de puntos cada uno de los cuales se corresponde con cada uno de los objetos dentro de la configuración.

Esta configuración se supone que refleja la estructura oculta de los datos, que al hacerse explícita a través del gráfico, nos facilita su comprensión más profunda.

Si se desea ver un ejemplo en el ámbito de la educación, se puede consultar el trabajo de Gil (1993), donde el autor explora el posicionamiento del profesorado respecto de la reforma educativa que se iniciaba en aquella época.

RESUMEN

La investigación ex post-facto agrupa un conjunto de métodos de gran aplicación en Educación. Este enfoque respeta el fenómeno tal como se produce en su ámbito natural y no pretende ni trata de manipular la realidad. Todo ello le otorga una capacidad ilimitada de actuar en ámbitos donde el ser humano es el centro y no sería ético actuar de otra manera a la par que no artificializa los fenómenos y recoge así la información en toda su riqueza y substantividad.

Bajo el epígrafe algo genérico y posiblemente excesivo de “ex post-facto”, hemos ido presentando un conjunto de estudios que hemos agrupado en los bloques siguientes:

Los estudios descriptivos, diseñados para aproximarnos a la realidad cuando desconocemos casi todo de ella y fundamentados en la descripción más rigurosa y sistemática de la misma.

Los estudios comparativo-causales, diseñados para actuar en ámbitos en que no es posible manipular las variables y se desea establecer relaciones de orientación causal.

Los estudios de desarrollo, en los que el tiempo constituye un factor primordial y nos interesa analizar su efecto en la evolución de los sujetos,

Los estudios correlacionales, en los que el proyecto fundamental consiste en explorar con toda la potencia de las técnicas multivariadas las complejas relaciones entre las variables más relevantes que intervienen en los fenómenos educativos.

Hemos de insistir en recordar la incapacidad formal de la investigación ex post-facto para establecer las relaciones causa-efecto, propias de la investigación experimental, pero también, y para finalizar, se ha de señalar la gran versatilidad y la enorme potencia de las técnicas que se aplican en estos modelos metodológicos que hacen de este tipo de investigación una de las opciones metodológicas con mayor futuro.

EJERCICIOS

1. Para realizar los siguientes ejercicios hay que leer previamente el documento “Modelos de investigación en la intervención educativa diferencial” (de M. Bartolomé y otros, 1994), publicado en la Revista de Investigación Educativa 23.
 - a) ¿Cuáles son los núcleos temáticos más amplios en los que se aplica la investigación ex post-facto?
 - b) ¿Cuáles son las críticas que hacen los autores respecto de los estudios que tratan de los efectos educativos en el rendimiento? Coméntalas brevemente.
 - c) ¿Qué tipos de estudio, desde el punto de vista metodológico, aplican Tokar y Swanson de un lado y Parker y Lord de otro?
 - d) En la página 39 del documento, en el punto 2.3, leemos en referencia con la investigación experimental: “El número de investigaciones experimentales es muy

inferior que el de los estudios ex post-facto, quizás por la dificultad real que comporta controlar las variables independientes. No hay demasiadas ocasiones en las que se pueda hacer la asignación aleatoria de los sujetos a los tratamientos". ¿Qué tiene que ver el control (o la falta de control) de la variable independiente y la preferencia de aplicar modelos ex post-facto?. Contesta y justifica brevemente tu respuesta.

2. Para realizar los siguientes ejercicios, es necesario leer el artículo de Blanca Rodríguez (1993), titulado "Modelo de evaluación externa de la eficacia de los centros educativos" y publicado por la Revista de Investigación Educativa en su número 21.
 - a) ¿Qué tipo de técnica multivariable ha hecho servir la investigadora de forma preferente?
 - b) ¿Qué programa informático ha utilizado para calcular los parámetros estructurales?
 - c) En el apartado 3.2 del artículo, la autora habla de variables exógenas y endógenas, que a su vez las subdivide en variables de proceso y de producto. Identifica cada una de ellas y confecciona una lista situándolas en el grupo correspondiente. ¿Cuál crees que es el criterio metodológico básico para clasificar una variable como exógena o endógena?

CAPÍTULO 7

ESTUDIOS DE ENCUESTA

Mercedes Torrado Fonseca

OBJETIVOS

- a) Conocer las características básicas de los estudios de encuesta.
- b) Saber identificar las diferencias de este tipo de estudios respecto al resto de modalidades de investigación de corte descriptivo.
- c) Identificar las fases propias del proceso de investigación mediante encuesta.
- d) Conocer e identificar las fortalezas y debilidades de este tipo de investigación.
- e) Saber seleccionar el tipo de encuesta más adecuado, a partir del objetivo de la investigación, del contexto, de los recursos disponibles, etc.
- f) Saber diseñar un proyecto de investigación mediante encuestas.

LOS ESTUDIOS DESCRIPTIVOS

A partir de diversos trabajos bibliométricos se constata que la práctica investigadora todavía hoy en día, tiende a utilizar metodologías de corte cuantitativo (Buendía, 2001). Aunque por otro lado, la necesidad de superación de la dicotomía cuantitativo *versus* cualitativo se hace cada vez más presente en los contextos científicos y prueba de ello son la amplia diversidad de investigaciones donde la pluralidad metodológica está presente.

Es evidente, pues, que estamos adentrándonos en una nueva etapa en la evolución de la investigación educativa donde aparece una superación de las cuestiones paradigmáticas. La necesidad de enfocar la investigación con una nueva perspectiva no debe

apartarnos de lo fundamental: *el objetivo y la finalidad de la investigación condicionará la modalidad de investigación más adecuada*, así como su metodología de investigación. Por este motivo los estudios de encuestas responden a una finalidad descriptiva específica y concreta. En muchas ocasiones corresponde a un primer acercamiento a la realidad que nos preocupa, para posteriormente poder estudiar con mayor profundidad el fenómeno educativo, utilizando otras modalidades de investigación más adecuadas al objetivo que se pretenda.

TABLA 7.1. *Relación entre objetivos de investigación no experimental y tipología de estudios.*

Objetivo que se pretende	Acciones a realizar	Tipo de estudio
Describir los fenómenos INVESTIGACIÓN DESCRIPTIVA	<i>Estudiar la evolución</i> ¿Existen cambios en los sujetos transcurrido un tiempo? <i>Recoger información</i> ¿Sabían los alumnos donde encontrar trabajo?, ¿Qué conductas de agresividad tienen los niños de 3 años?	Estudios de desarrollo Longitudinales transversales cohortes Estudios de encuestas Estudios observacionales
Relacionar y predecir valores de fenómenos INVESTIGACIÓN CORRELACIONAL	<i>Relacionar</i> ¿En qué condiciones se produce un fenómeno? ¿cuál es el grado de probabilidad de que suceda? <i>Predecir</i> ¿Cuántos alumnos abandonarán los estudios universitarios?	Estudios correlacionales Estudios predictivos
Explicar relaciones de causalidad entre fenómenos INVESTIGACIÓN COMPARATIVO-CAUSAL	<i>Comparar</i> ¿Qué método didáctico influye más en el rendimiento académico?	Estudios comparativo-causales

Para entender mejor los estudios por encuesta, es importante partir de una concepción clara de las investigaciones descriptivas. Por ello, creemos pertinente presentar a modo de resumen aspectos clave de la investigación descriptiva.

La investigación descriptiva tiene como objetivo general describir un fenómeno dado, a partir de diversas acciones no excluyentes entre ellas. Tal y como manifiesta Mateo (1997) este tipo de estudio es propio de las primeras etapas del desarrollo de una investigación y pueden proporcionar datos que vayan preparando el camino para nuevas investigaciones.

Respecto a la clasificación de la investigación descriptiva no existe un acuerdo unánime entre la comunidad científica. Partimos, pues, de nuestra propia clasificación. Así, nosotros diferenciamos tres tipos de estudios descriptivos: estudios de desarrollo, estudios de encuestas y estudios observacionales.

En la siguiente tabla se presenta la clasificación de las investigaciones de orientación cuantitativa, pero no experimentales, donde se ubican los estudios de encuesta. Esto facilita situarnos mejor en el complejo mundo de la investigación educativa.

DEFINICIÓN DE LOS ESTUDIOS DE ENCUESTA

La investigación por encuesta es muy utilizada en el ámbito educativo, quizás debido a su aparente facilidad. Incluye los estudios que utilizan como instrumento de recogida de datos el cuestionario y los que se llevan a cabo mediante entrevistas. Son útiles para la descripción y la predicción de un fenómeno educativo, pero también son eficientes para una primera aproximación a la realidad o para estudios exploratorios.

Pero ¿qué entendemos y cómo definimos los estudios de encuesta. Son muchas las definiciones que van apareciendo en la bibliografía básica de investigación educativa. No creemos oportuno presentar un listado exhaustivo de todas ellas, pero sí cuáles son los elementos o rasgos propios de los estudios de encuesta, independientemente del uso del instrumento de recogida de datos, que nos ayudarán a definir este tipo de investigación descriptiva. Así, los estudios de encuesta permiten:

- Recoger información de los sujetos a partir de la formulación de preguntas a través de una entrevista personal, por correo o por teléfono;

- Y pretender hacer estimaciones de las conclusiones a la población de referencia a partir de los resultados obtenidos en una muestra.

En la práctica habitual de los estudios de encuestas podemos encontrar ejemplos donde la selección de una muestra no sea el elemento clave de la investigación, sino la descripción de la totalidad de la población, sea una clase, escuela, barrio, etc. También podemos encontrar investigaciones que utilizan diversos procedimientos de recogida de información.

INTRODUCCIÓN HISTÓRICA A LOS ESTUDIOS POR ENCUESTA

La aplicación y uso de los estudios de encuesta no responden a ninguna moda ni interés político, todo lo contrario. Su origen y aplicación en el ámbito educativo está relacionado con el nacimiento de la investigación educativa como disciplina. Sin embargo, sus antecedentes son anteriores. Basándonos en la obra de Arias y Fernández (1998) expondremos en este apartado una síntesis de los diferentes movimientos más cercanos al inicio de la investigación a partir de encuestas: la Aritmética Política, la Estadística Moral, la Medición de las actitudes y los Censos electorales. Tal y como recoge Arias y Fernández (1998) las encuestas son una consecuencia clara de uno de los *dictums* más importantes de la historia social de la modernidad formulados por Comte: “*conocer para prever y el prever para poder*”.

La idea de que es posible acceder al conocimiento de la realidad mediante el análisis cuantitativo de los hechos que la conforman tiene su primera e importante defensa en la Inglaterra de la segunda mitad del siglo XVII. Así, Bacon, siguiendo la convicción dominante del momento, crea el clima adecuado para el desarrollo de importantes metodologías; en concreto, destaca lo que se conoce con el nombre de *Aritmética Política*. Graunt, en 1662, sirviéndose de registros parroquiales, elabora diversas tasas de natalidad y mortalidad de la población de Londres, publicadas en su obra *Natural and political observations made upon the bills of mortality*. Se puede decir que fue el primero en razonar de manera estadística sobre material demográfico. Como continuación en esta línea, destacamos el trabajo de Gregory King en Inglaterra, *Natural and Political Observations upon the state and condition of*

England (1916), en el cual, además de estimar el número de familias del país, calculaba la media de ingresos familiares.

Ya en el siglo XIX aparece otra tradición relacionada con la investigación mediante encuestas. Estamos hablando de la *Estadística Moral*. Autores como el ingeniero de minas Leplay (francés) y Quetelet (belga) suelen estar asociados a este movimiento. En este caso las aportaciones de Quetelet están relacionadas directamente con la utilización de tablas multivariadas para explorar relaciones entre diferentes tasas, así como la aplicación del cálculo de probabilidades para explicar la constancia en el tiempo de determinados fenómenos.

A principios del siglo XIX se funda la *Sociedad Estadística* de Inglaterra. A partir de este momento aparece una preocupación por obtener datos sobre temas como la salud, la educación, etc., apareciendo el movimiento de las *Encuestas y Monografías Sociales* que se extiende rápidamente por Europa y los Estados Unidos. De este movimiento destacamos la aportación de Emile Durkheim, ya que basándose en los datos acumulados por la Estadística Moral ofrece una explicación sociológica del suicidio.

No podemos obviar en este recorrido histórico el auge de la *medición de las actitudes*. Los primeros antecedentes aparecen en 1817 con Marc Antoine Jullien de París, quien diseñó un cuestionario nacional e internacional sobre todos los aspectos relacionados con el sistema educativo. En 1892 Rice realiza casi 1.200 entrevistas a profesores sobre el currículum y los métodos de enseñanza. En 1904 Thorndike publica un manual sobre métodos de investigación asociando la información científica con la medida; poco más tarde, en 1907, elabora un cuestionario sobre el abandono escolar y, posteriormente, realiza una serie de encuestas escolares sobre las diferencias curriculares, el fracaso escolar, la calidad del profesorado, el equipamiento escolar, etc. Thurstone en 1928 y Likert en 1932 publican sendos artículos en los que presentaban técnicas muy precisas y elaboradas a partir de las cuales se demostraba que era posible medir las actitudes.

Por último, encontramos los *Resultados Electorales* como el antecedente más importante de las encuestas electorales. De este movimiento destacamos el estudio realizado en 1936 por la revista *Literary Digest* para predecir con éxito los resultados electorales.

De esta forma se generalizó el uso de muestras. Hoy en día las encuestas de intención de voto aparecen frecuentemente en la prensa.

PROCESO DE INVESTIGACIÓN

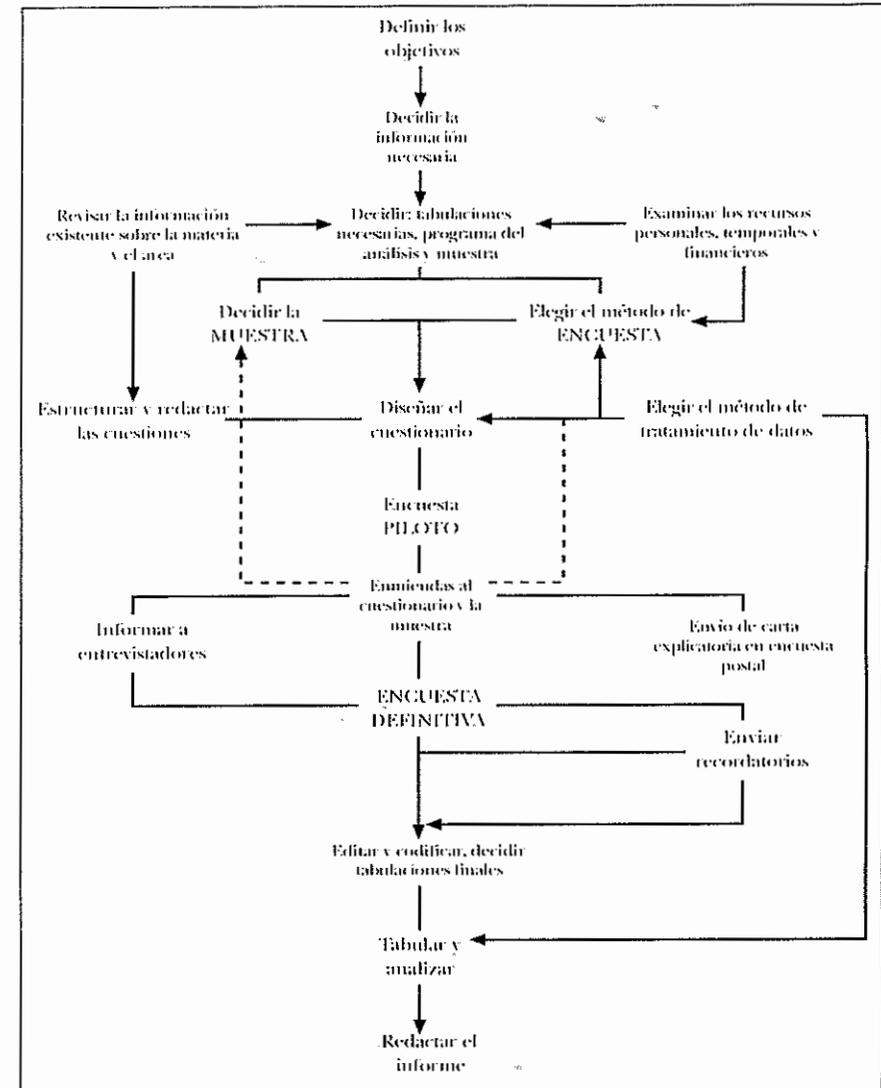
En el proceso de investigación existen varias etapas. Cohen y Manion (2002) plantean uno de los esquemas más completos en cuanto al proceso de investigación por encuesta, que ha servido de base para otros esquemas más sintéticos. Así, Buendía (1998) establece tres fases de desarrollo: teórico-conceptual, metodológica y estadístico-conceptual. En el primer nivel incluye el planteamiento de los objetivos y/o problemas e hipótesis de investigación. En el segundo se ubica la selección de la muestra y la definición de las variables que van a ser objeto de estudio. La confección del cuestionario piloto y su formulación definitiva darán paso al nivel estadístico, en el que, tras la codificación y análisis de los datos, se podrán elaborar conclusiones, realizar generalizaciones e integrar en el marco teórico de partida las conclusiones elaboradas. Por su parte, Arnau (1995) plantea cuatro etapas en el proceso de investigación por encuesta, la primera relacionada con el planteamiento de objetivos y preparación del instrumento de recogida de información, seguido de la planificación del muestreo, recogida de datos y, finalmente, análisis e interpretación de los datos.

Partiendo de los planteamientos anteriores, en todo proceso de investigación, independientemente de la modalidad, debe quedar muy claro cuáles son los objetivos del estudio. Los objetivos deben llevar fácilmente a establecer cuál es la información necesaria y pertinente que debe recoger el instrumento de recogida de datos.

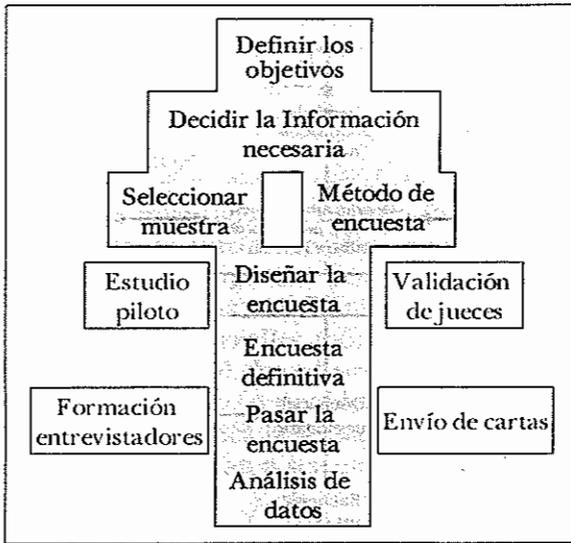
Selección de la muestra

La imposibilidad, en muchos de casos, de no poder trabajar con todos los sujetos de la población obliga a tener que elegir una muestra que represente en todas sus características y peculiaridades a la población de referencia.

El hecho de trabajar con un número reducido de sujetos conlleva una serie de ventajas que se deben considerar. Por ejemplo,



Esquema 7.1. Proceso de investigación por encuesta (Cohen y Manion, 2002: 133).

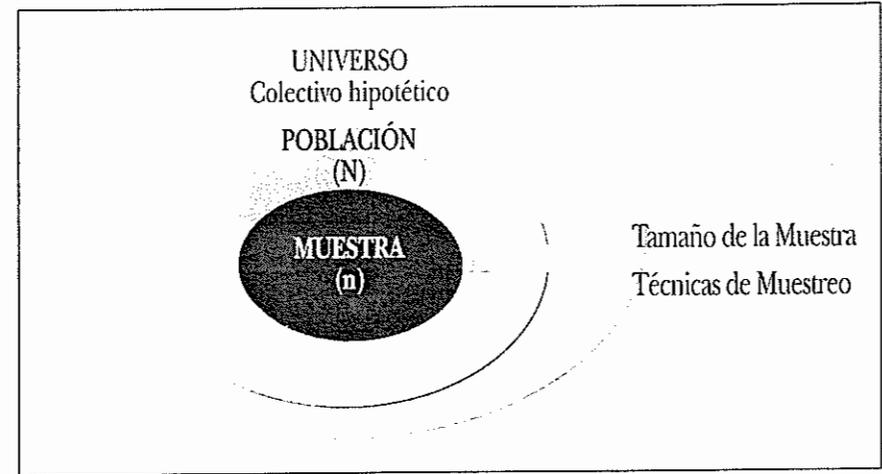


Esquema 7.2. Fases de la investigación por encuesta.

presenta una reducción importante de costos. Pero, a su vez, también exige una rigurosidad en el proceso de selección. A continuación presentamos los aspectos más relevantes.

a) *Representatividad de la muestra*

Incidiremos en destacar sobre la importancia de trabajar con una muestra representativa en este tipo de estudio. En primer lugar, debemos detectar aquellas variables que posibiliten una descripción de las características propias de la población y que deberán estar presentes en la muestra seleccionada. Las mencionadas variables, denominadas *variables marcadoras* (Martínez Arias, 1995) o simplemente variables identificadoras de la población, corresponden a factores demográficos de los sujetos: sexo, edad, estatus socioeconómico, tipo de escuelas,... Hay que insistir en la utilización de *técnicas de muestreo probabilísticas* como mecanismo para lograr el objetivo de la representatividad.

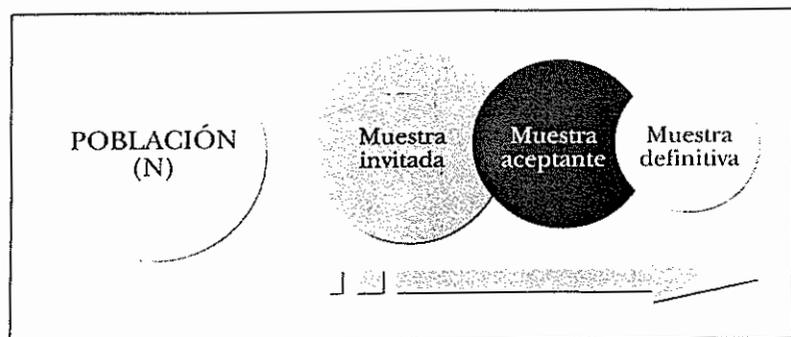


Esquema 7.3. Selección de la muestra.

Además de la técnica de selección de los sujetos de la muestra previamente se ha tenido que determinar el número mínimo de sujetos: *¿Cuántos sujetos debe tener la muestra para ser representativa?* El tamaño de la muestra dependerá del propósito de la investigación y de la población objeto; por lo tanto, si se pretende realizar diversas pruebas estadísticas, se deberá trabajar con un número suficiente de sujetos. A grandes rasgos el número mínimo dependerá del número de sujetos que comprende la población, de la magnitud de heterogeneidad de la variable, del nivel de confianza y del error máximo con el que se decida realizar el estudio. También es importante introducir el concepto de error de muestreo o error muestral. Existen fórmulas estadísticas para el cálculo del número mínimo de sujetos.

b) *El ciclo del muestreo. Pérdida de sujetos*

Muy a nuestro pesar, la muestra seleccionada no coincide con la muestra final o productora de datos. En algunos casos esta diferencia puede invalidar todo el proceso de selección de la muestra. Por ello y siguiendo a Fox (1981) presentamos el ciclo de muestreo.



Esquema 7.4. Ciclo de muestreo.

TÉCNICAS DE RECOGIDA DE INFORMACIÓN

La investigación por encuesta debe traducir las variables sobre las que se desea obtener información en preguntas concretas sobre la realidad. Las técnicas asociadas a este tipo de investigación son el cuestionario y la entrevista, utilizándose en muchas ocasiones como sinónimos. Pero no cabe duda que los distintos sistemas de recopilación de la información implican diferentes grados de interacción entre el entrevistador y el entrevistado a la hora de presentar la encuesta.

- Un *cuestionario* es un instrumento de recopilación de información compuesto de un conjunto limitado de preguntas mediante el cual el sujeto proporciona información sobre sí mismo y/o sobre su entorno.
- Una *entrevista* es un procedimiento mediante el cual un entrevistador realiza un conjunto de preguntas a un sujeto. Las preguntas pueden estar totalmente definidas de forma previa (entrevista estructurada) o bien estar indefinidas en menor o mayor grado (entrevista semiestructurada).

Tenemos que considerar que la elaboración de la encuesta no empieza con la redacción de las preguntas, sino que requiere una serie de fases:

1. Definición de los objetivos del cuestionario.
2. Planificación del cuestionario, destacando sus diferentes apartados.
3. Elaboración y selección de las preguntas (abiertas, cerradas, ...).
4. Análisis de la calidad de las preguntas.
5. Análisis de la fiabilidad y validez del cuestionario.
6. Redacción final de la encuesta.

CLASIFICACIÓN DE LOS ESTUDIOS DE ENCUESTA

Existen múltiples modos de clasificar las encuestas: en función de sus características, según el grado de cobertura de la población, en función de su objetivo, de su diseño, de la técnica de muestreo, etc. En este apartado se exponen diversos criterios clasificatorios.

TIPO DE ENCUESTAS EN FUNCIÓN DE LA FORMA

Se suelen distinguir los tipos de encuestas en función de la forma en la que se administran. A continuación presentaremos las ventajas y desventajas de las entrevistas personales, entrevistas o encuestas por teléfono, las encuestas postales y, por último, las encuestas por internet.

La *entrevista personal* es el tipo de encuesta que con mayor frecuencia se utiliza. Implica la participación directa del entrevistador que es quien plantea las cuestiones a los sujetos entrevistados. En este caso, el entrevistador deberá cumplir una serie de requisitos para que las respuestas puedan ser comparadas y analizadas con cierta fiabilidad. Por ello deberán ser:

- Persona con cierta formación y experiencia en la realización de entrevistas personales.
- Deberá formular en el mismo orden y con los mismos planteamientos las diferentes cuestiones de la entrevista.
- Deberá utilizar un guión de entrevista o cuestionario.

TABLA 7.2. Entrevistas personales.

ENTREVISTAS PERSONALES	
Fortalezas	Debilidades
<ul style="list-style-type: none"> • Facilita la cooperación con los sujetos y permite establecer con ellos una relación de confianza. • Permite aclarar posibles dudas en cuanto al redactado de preguntas o seguimiento de las instrucciones. • Suele proporcionar elevadas tasas de respuesta. 	<ul style="list-style-type: none"> • El entrevistador es una fuente clara de invalidez, tanto por cómo plantea las preguntas, por sus características personales, etc. • Requiere una formación y entrenamiento del instrumento de medida por parte del grupo de entrevistadores. • Costos elevados. • Duración temporal elevada.

Las *encuestas telefónicas* presentan algunos rasgos comunes con las entrevistas personales, sobre todo en lo relacionado con los entrevistadores.

TABLA 7.3. Entrevistas telefónicas.

ENTREVISTAS TELEFÓNICAS	
Fortalezas	Debilidades
<ul style="list-style-type: none"> • Menor coste que las entrevistas personales. • Permiten acceder a las poblaciones de difícil acceso. • La selección de la muestra es más fácil. • El proceso de recogida de datos es relativamente más rápido. • Menor número de entrevistadores que en las entrevistas personales. • Permite aclarar dudas en las formulaciones de preguntas. • Tasas de respuestas elevadas. 	<ul style="list-style-type: none"> • Problemas de contacto con algunas personas seleccionadas (teléfonos incorrectos o cambios de domicilio, ...). • Imposibilidad de plantear preguntas complejas. • Procedimiento no recomendado para el planteamiento de temas delicados y sensibles. • Menor cooperación de los sujetos • Formación y entrenamiento del equipo de entrevistadores.

Las *encuestas postales* implican la utilización del correo como mecanismo para hacer llegar la encuesta a una muestra representativa. A diferencia de las dos anteriores, este tipo de encuesta presenta dos rasgos propios. Por un lado, el instrumento de auto-cumplimentación y, consecuentemente, las preguntas e instrucciones de respuesta deben ser entendibles. Este hecho conlleva realizar convenientemente una prueba piloto previa al pase definitivo, así como validar el contenido de las cuestiones planteadas mediante el proceso de *validación de jueces*. Si bien esto es conveniente también en todo tipo de encuesta.

Validación del instrumento: Prueba Piloto

En las encuestas con grandes poblaciones suele ser un requisito imprescindible realizar una prueba piloto, que consiste en la aplicación previa en menor escala de todos los procedimientos que se utilizarán en la encuesta final. Para la realización de la prueba piloto no suele ser necesario utilizar muestras representativas, por lo que suelen emplearse técnicas de muestreo no probabilísticos. A veces es suficiente con muy pocos sujetos. A partir de esta experiencia previa se podrán realizar estudios de fiabilidad y validez del instrumento mediante una serie de pruebas estadísticas.

Validación de jueces

Se seleccionan un grupo de personas expertas en la materia que se pretende estudiar y se les plantea una serie de cuestiones en cuanto al contenido del cuestionario y a su estructura. Por ejemplo: *¿Las preguntas están bien formuladas? ¿El contenido de las preguntas engloba todas las dimensiones del fenómeno? ¿Qué aspecto consideras que pretende estudiar cada una de las preguntas? ¿A qué dimensión conceptual se adscriben las cuestiones? etc.*

TABLA 7.4. Encuestas postales.

ENCUESTAS POSTALES	
Fortalezas	Debilidades
<ul style="list-style-type: none"> • Bajo coste, hecho que posibilita incluir más sujetos. • No se necesitan entrevistadores, evitando los posibles sesgos del entrevistador. • Al entrevistado le produce mayor sensación de anonimato. • Permite acceder a las poblaciones de difícil acceso. • Permite al entrevistado responder con calma a determinadas cuestiones. 	<ul style="list-style-type: none"> • Menores tasas de respuesta. • Los entrevistados que no responden suelen poseer unas características diferentes de los que sí lo hacen. • Ante preguntas sin responder no existe ninguna posibilidad de poder cumplimentarla. • Las instrucciones deben ser muy claras y sencillas. • No permite observar reacciones de los sujetos.

Se puede afirmar que la encuesta por correo es la estrategia con menor tasa de respuesta. Existen varias estrategias para subsanar esta debilidad. En algunas ocasiones el seguimiento de los que no responden por medio de entrevista telefónica puede ser una solución que aumenta la tasa de respuesta de forma inmediata. Otra forma es enviar nuevamente una encuesta más corta y por lo tanto de mayor facilidad en su cumplimentación. Si el porcentaje de respuesta espontáneo oscila alrededor del 40%, cuando utilizamos algún tipo de refuerzo puede aumentar hasta un 60% (Martínez Arias, 1995) o incluso entre un 70% y un 80% (Cohen y Manion, 2002).

La *entrevista por internet* es cada vez es más habitual. Tiene las mismas características que las anteriores, excepto que el medio es internet.

Criterios para la selección del procedimiento de recogida de información

Hasta ahora hemos planteado las ventajas y los inconvenientes en cada uno de los tipos de encuesta o entrevista, pero ¿cuáles son los criterios generales que debemos considerar en el momento de la selección del procedimiento más adecuado?

TABLA 7.5. Encuestas por internet.

ENCUESTAS POR INTERNET	
Fortalezas	Debilidades
<ul style="list-style-type: none"> • Bajo coste y mayor rapidez. • No se necesitan entrevistadores, evitando los posibles sesgos del entrevistador. • Al entrevistado le produce mayor sensación de anonimato. • Permite acceder a las poblaciones de difícil acceso. • Evita errores de codificación, saltos de preguntas, etc. • Facilita el análisis cuantitativo de la información recogida. 	<ul style="list-style-type: none"> • Dificultad en plantear preguntas abiertas. • Mayor coste en la elaboración del cuestionario. • Riesgos derivados de posibles fallos informáticos. • Necesidad de disponer conexión a internet. • Es un medio poco motivador si no existe un reclamo o incentivo. • Es muy impersonal. • Imposibilidad de encuestar a quienes no dispongan de esta tecnología.

A partir de una serie de criterios, como versatilidad, coste, control de la muestra, etc., hemos confeccionado la siguiente tabla resumen.

TABLA 7.6. Criterios de selección del tipo de encuesta.

Criterios	Entrevista personal	Entrevista telefónica	Entrevista Correo/internet
Versatilidad (Capacidad de adaptación)	+	+/-	-
Coste (Número de horas de trabajo)	+	+/-	-
Duración del período de recogida de información	+	-	+/-
Control de la muestra	+	+	-
Cantidad de información	+	-	+/-
Calidad de la información	+	+	+/-
Tasa de respuesta	+	+	-

Consideraciones generales sobre la elaboración del instrumento

En el momento de elaborar el cuestionario es necesario saber cuál es la información exacta que queremos recoger. Por ello, en muchas ocasiones antes de proceder a la confección de las diversas preguntas se utilizan o bien encuestas abiertas o se trabaja a partir de grupos de discusión.

El hecho de elaborar mapas conceptuales previos también es una estrategia utilizada en la fase de elaboración del instrumento. De esta forma, quedan reflejados todos los aspectos relevantes que debe contemplar la encuesta.

TABLA 7.7: Comparación de preguntas abiertas y cerradas (Padilla, González y Pérez, 1998: 133).

	<i>Preguntas ABIERTAS</i>	<i>Preguntas CERRADAS</i>
<i>Objetivo</i>	Obtener información directa de los encuestados con sus propias palabras	Obtener datos de juicios u ordenaciones
<i>Características de los encuestados</i>	Nivel cultural medio-elevado y con conocimiento y opinión sobre el tema	Nivel cultural medio-bajo y sin conocimiento y opinión sobre el tema
<i>Forma de hacer las preguntas</i>	No disponer de las posibles respuestas. Posibles respuestas muy numerosas	Respuestas previstas
<i>Análisis de los resultados</i>	Posibilidad de categorizar y analizar respuestas muy heterogéneas	Descripción meramente cuantitativa de las respuestas
<i>Comunicar los resultados</i>	Descripción individual de los tipos y categorías de respuesta	Información estadística de las respuestas

En un cuestionario conviene introducir:

- Variables de identificación y de clasificación, como puedan ser género, edad, nivel de estudios, etc...

- Las preguntas abiertas deberán ser concisas y de respuesta breve, ya que de lo contrario el encuestado no suele responderlas.
- Considerar el orden de las cuestiones. Presentar las preguntas más interesantes al principio. También agruparlas por temáticas para evitar la desorientación en el momento de rellenar o responder.
- El cuestionario o entrevista no debe ser excesivamente larga.
- La redacción de las preguntas debe ser sencilla, utilizando un lenguaje apropiado para las personas a las que va dirigido el instrumento.
- El aspecto formal debe considerarse, sobre todo en las encuestas por correo.
- Se debe redactar una *presentación* que especifique el objetivo de la investigación y que se pretende.

Especial atención a las encuestas de correo

Las encuestas por correo es un tipo de cuestionario de auto-cumplimentación y por lo tanto tiene que tener unas características de presentación especiales que van más allá del contenido del mismo.

- *El aspecto del cuestionario.* Debe parecer fácil y atractivo de responder. Las preguntas no deben estar muy apretadas, sino presentar espacio suficiente para responder. También en algunas ocasiones se utiliza una impresión en color, diferenciando partes del cuestionario.
- *La redacción debe ser clara,* tanto en cuanto a las preguntas como en las mismas instrucciones utilizando frases como “*marque con una cruz*”, “*marque su grado de acuerdo en la escala de valoración*”.
- *La distribución del contenido* debe facilitar su cumplimentación. Por lo tanto, las cuestiones referentes al mismo tema deben estar juntas y no distribuidas a lo largo del cuestionario. Considerar y controlar el posible aburrimiento y la frustración.
 - Las preguntas iniciales deben ser sencillas, tener interés y animar a la participación.

- En la sección central se deberán concentrar las preguntas difíciles.
 - Mantener el interés por devolver el cuestionario.
- El final del cuestionario suele dedicarse para rogar al encuestado que repase toda la información, solicitar una devolución rápida y sobre todo agradecer la participación y colaboración: "*Gracias por su colaboración desinteresada*".

Junto con el cuestionario debemos incluir una carta de presentación donde se presente claramente el objetivo de la encuesta y comunicar su importancia. También se debe poner de manifiesto la garantía de la confidencialidad e informar sobre cómo se han accedido a sus datos personales. Por último se debe facilitar el nombre y teléfono del remitente para que el encuestado pueda en un momento determinado consultar la veracidad del estudio. Para facilitar su respuesta se deberá adjuntar un sobre de retorno con la dirección y el franqueo correspondiente.

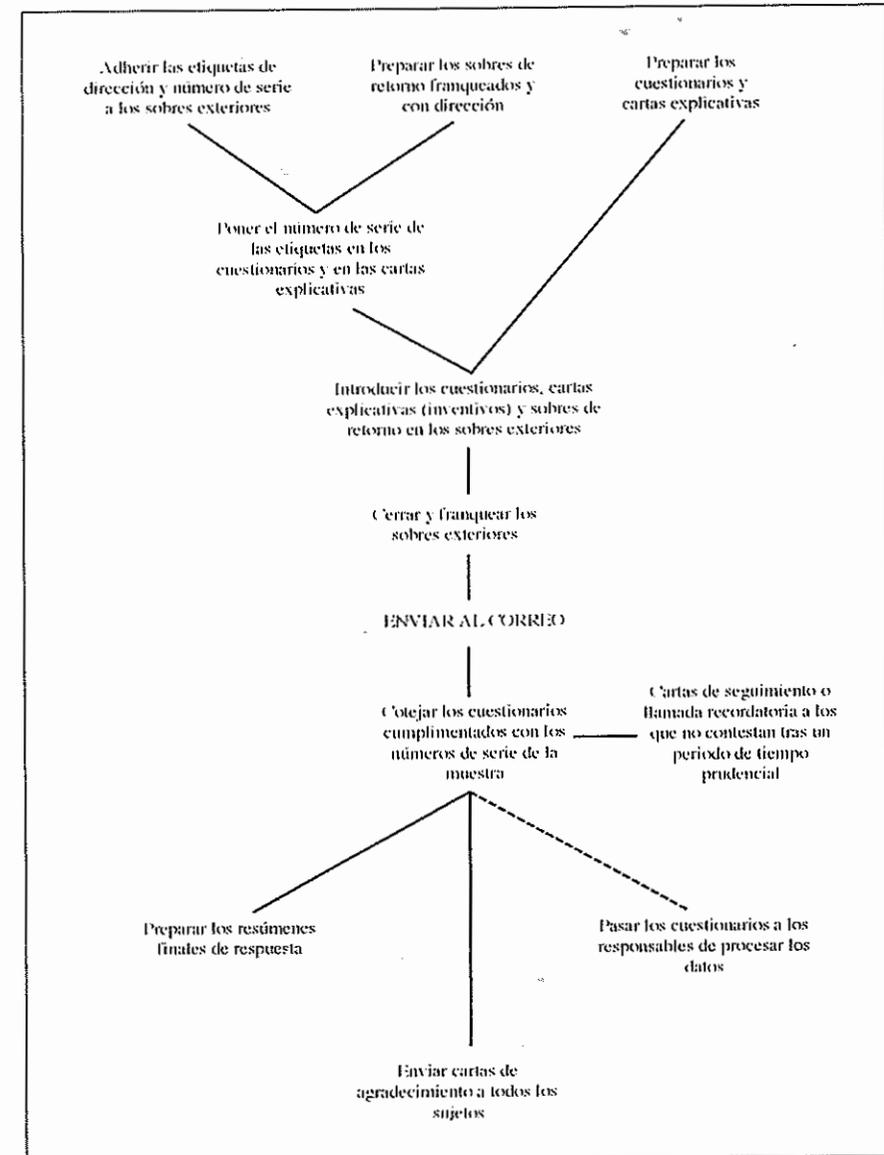
Aparte de la calidad de los sobres y ofrecer una imagen de seriedad, debe considerarse el día del envío y, lo que es más importante, la época. Evitaremos períodos de vacaciones con el fin de garantizar la recepción del envío.

Dado que la encuesta por correo es el procedimiento con menor tasa de respuesta, existen mecanismos para aumentarla. Básicamente son dos:

- a) Una carta de seguimiento al cabo de dos semanas del primer envío, en la cual se adjunte nuevamente el cuestionario.
- b) Un recordatorio telefónico, confirmando la recepción del cuestionario, aprovechando la clarificación de las dudas del mismo e incidiendo sobre la importancia y relevancia de su devolución.

Otro factor que hay que considerar en este tipo de encuesta es el uso de incentivos, sobre todo en aquellos casos en que fuese procedente un segundo o tercer contacto por constituir un estudio longitudinal.

Esquema 7.5: Diagrama de flujo para planificar una encuesta por correo (Cohen y Manion, 2002: 152).



Tipos de encuestas según la finalidad

Una posible clasificación de los estudios de encuesta en función de los objetivos es la siguiente.

TABLA 7.8. *Tipos de estudios de encuestas según la finalidad.*

Clasificación según	TIPOS DE ESTUDIOS DE ENCUESTAS
Objetivo	DESCRIPTIVAS EXPLICATIVAS
Dimensión temporal	TRANSVERSALES LONGITUDINALES
Sentido temporal	RETROSPECTIVAS PROSPECTIVAS

Si la finalidad u *objetivo de la investigación* es conocer o describir las características de la población, a partir de un estudio detallado de las variables de interés tendremos que elaborar una encuesta descriptiva. En cambio, si la finalidad de la investigación es establecer relaciones causales y/o contrastar hipótesis se planificarán encuestas explicativas.

En la siguiente tabla se recogen cuáles son las implicaciones propias de las encuestas descriptivas y explicativas.

TABLA 7.9. *Encuestas descriptivas y explicativas.*

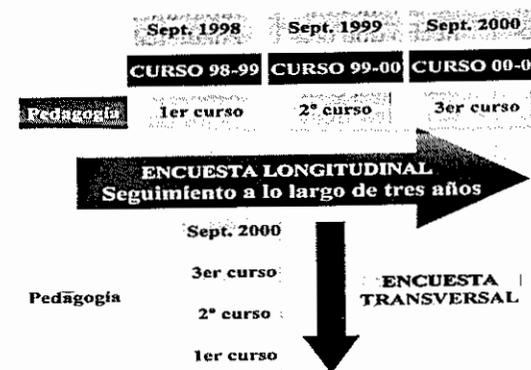
ENCUESTAS DESCRIPTIVAS	ENCUESTAS EXPLICATIVAS
<ul style="list-style-type: none"> • Generalmente al inicio del instrumento aparecen datos personales que permitan describir la muestra y también realizar comparaciones entre las diferentes categorías. • Se trabaja con muestras representativas con el objetivo de poder realizar inferencias a la población de origen. 	<ul style="list-style-type: none"> • Normalmente se ha realizado un estudio exploratorio previo (encuestas descriptivas). • Los sujetos de la muestra deberán poseer valores de la variable independiente, debido a la imposibilidad de una manipulación por parte del investigador. • Utiliza técnicas estadísticas que permiten confirmar relaciones causales.

Hay que señalar, no obstante, que la diferencia entre encuestas descriptivas y encuestas explicativas es más teórica que real; en muchas ocasiones no son mutuamente excluyentes sino que, por el contrario, se encuentran habitualmente combinadas en la misma encuesta.

Al abordar la clasificación según la *dimensión temporal* también estamos presentado algunos de los estudios de desarrollo. Así pues, distinguimos entre los estudios transversales y los estudios longitudinales. Los primeros, caracterizados por recoger información una única vez en un período de tiempo delimitado y de una población definida, son considerados como especialmente adecuados para el estudio de variables estables en el tiempo. En los segundos se trata de seguir la evolución de un sujeto o un grupo durante un periodo de tiempo.

El siguiente esquema ilustra gráficamente cuáles son las diferencias entre los estudios longitudinales y los transversales en cuanto a diseño y planificación. Partimos de dos investigaciones con un mismo objetivo: *estudiar la transición de los estudiantes universitarios de Pedagogía en sus tres primeros años de carrera*. En el primer caso se seleccionan todos los estudiantes que se matricularon en los estudios de Pedagogía en el curso 1998-99 y se realiza un seguimiento de los mismos en los cursos sucesivos. En cambio, en el segundo caso, se seleccionan estudiantes de Pedagogía de cursos diferentes en un mismo momento.

Esquema 7.6. Estudios longitudinales y transversales.



Pero, ¿cuáles son las ventajas y desventajas de unas u otras encuestas?

Las encuestas longitudinales tienen como objetivo recoger información a lo largo de un período de tiempo, llegando, en algunas ocasiones, a plantear estudios de varios años. Cuando se dirigen a los mismos sujetos se pueden denominar estudios de seguimiento o *estudios de grupo o de cohortes*, término equivalente en los Estados Unidos como *estudios de panel* (Cohen y Manion, 2002). Si se estudian diversos factores continuamente en el tiempo se denominan estudio de tendencias.

Sobre todo en las encuestas longitudinales el instrumento debe ser similar para poder estudiar la evolución o tendencia de los sujetos. También debe mantener las mismas convenciones de codificación y edición de los datos. A modo de resumen la siguiente tabla presenta las fortalezas y las debilidades de cada una de ellas.

Tabla 7.10. *Estudios longitudinales y transversales.*

ESTUDIOS LONGITUDINALES		ESTUDIOS TRANSVERSALES	
Fortalezas	Debilidades	Fortalezas	Debilidades
<ul style="list-style-type: none"> • Con muestras representativas, son estudios útiles para identificar esquemas típicos de desarrollo y para identificar factores. • Permiten identificar variaciones individuales a partir de registros longitudinales. • Son apropiados cuando se intenta establecer relaciones causales 	<ul style="list-style-type: none"> • Son estudios que requieren mucho tiempo, resultando caros. • Son estudios afectados inevitablemente por la mortalidad de la muestra, hecho que ocasiona, en algunos casos, escasa representatividad. • La repetición de entrevistas ocasiona efectos indeseados como el efecto control o efecto medición. El sujeto se sensibiliza. 	<ul style="list-style-type: none"> • Son estudios más económicos. Hecho que ocasiona poder trabajar con más sujetos. • Producen resultados más rápidamente. • Puede haber efectos indeseados provocados por la medición. • No existe mortalidad de la muestra. 	<ul style="list-style-type: none"> • Es un método menos eficaz para identificar variaciones individuales y poder establecer relaciones causales. • La selección de los sujetos es más complicada ya que están involucrados sujetos de diferentes niveles de edades y pudieran no ser comparables.

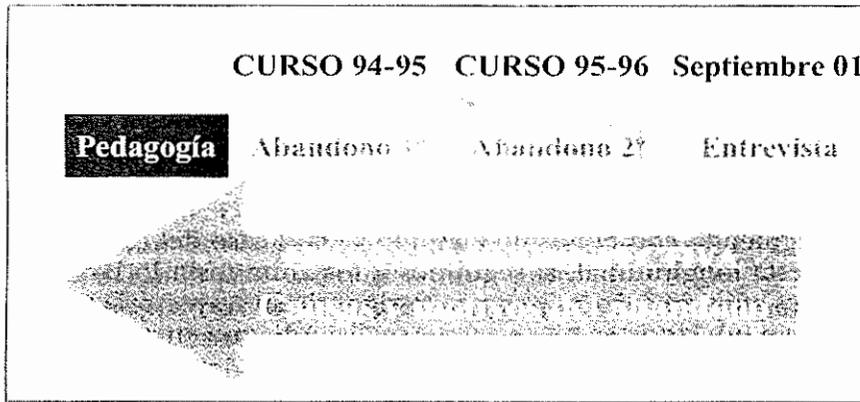
Según el *sentido y momento de recogida* de información nos podemos plantear estudios retrospectivos o prospectivos. Son prospectivos los estudios de tendencias que avanzan en la recopilación de los datos, a la vez que se van sucediendo los hechos. Sin embargo, cuando se pretende estudiar la situación actual a partir de sucesos acontecidos en el pasado estamos planteando un estudio retrospectivo.

Siguiendo con el mismo ejemplo y en el caso del estudio de encuesta longitudinal, si seguimos a los alumnos a lo largo de los tres años de carrera y somos testigos de su trayectoria académica sin interferir en ella estamos realizando una encuesta prospectiva.



Esquema 7.7. Encuesta prospectiva.

Sin embargo, si partimos del presente y pretendemos estudiar qué aconteció en los dos primeros años de carrera de un conjunto de estudiantes que abandonaron sus estudios, para poder describir e identificar posibles causas y motivos del abandono o deserción universitaria, nuestra mirada es hacia el pasado y no hacia el futuro. Estamos haciendo un estudio retrospectivo.



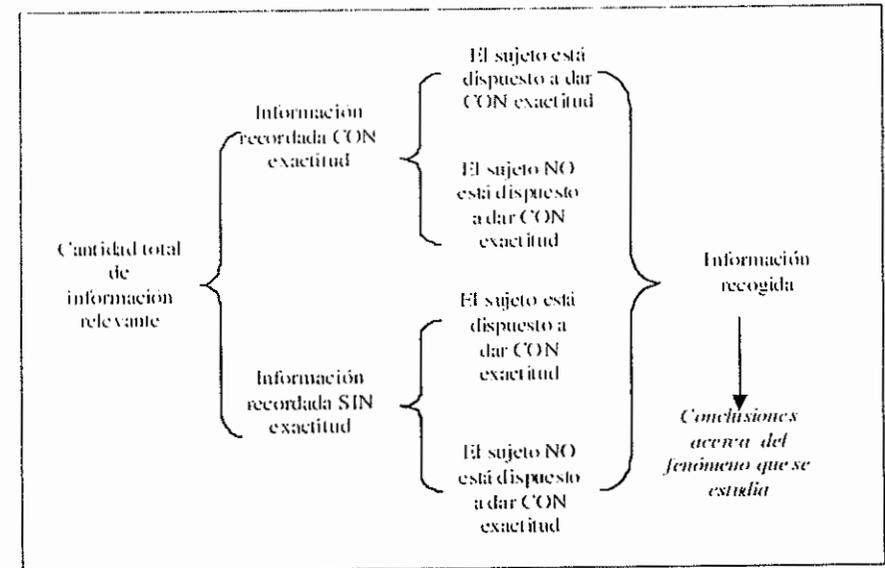
Esquema 7.8. Encuesta retrospectiva.

Dentro de las dos modalidades que presentan los estudios longitudinales, el retrospectivo¹ es el más utilizado frente al prospectivo. Pero, independiente de la modalidad de análisis, este hecho no elimina los graves problemas que presentan este tipo de investigación destacando dos principalmente: la pérdida de sujetos y la maduración personal. Estos efectos van aumentando, como fuentes de invalidez interna, a medida que la investigación se prolonga en el tiempo.

Otra limitación que presentan los estudios retrospectivos, si la fuente de información es el propio sujeto, es la dificultad del tiempo. Por un lado, en cuanto al problema que supone localizar a los sujetos después de un período de tiempo largo y, por otro, a la validez de la información que se recoge a través del propio sujeto por su dificultad de confirmación. El siguiente esquema, adaptado de Fox (1981), ilustra perfectamente la debilidad de este tipo de investigación en cuanto a la cantidad y calidad de la información recogida a través de la encuesta o entrevista.

En los estudios retrospectivos existen dos tipos de información, la que se recuerda con exactitud y la inexacta, que están for-

maadas a su vez por dos subpoblaciones, según la información que el sujeto está dispuesto a facilitar y la que está dispuesto a dar de forma conscientemente distorsionada. Con lo cual es una limitación importante en este tipo de estudios: el sujeto consultado puede tratar de crear una descripción del pasado favorable o desear modificar la imagen de lo que hizo en el pasado. A esto se le denomina "efecto de deseabilidad social".



Esquema 7.9. Calidad y cantidad de información a partir de una muestra retrospectiva (Fox, 1981: 186).

Ahora bien, si los datos no se extraen de los recuerdos del sujeto, sino más bien, de fuentes documentales como bases de datos, este tipo de investigación adquiere mayor fortaleza, ya que dichos datos no han sido modificados ni afectados por agentes externos. A modo de resumen la siguiente tabla recoge las ventajas y desventajas propias de los estudios retrospectivos por encuestas.

¹ Kerlinger denomina a este tipo de estudio ex post-facto y Dale y Meyer estudios "comparativo-causales".

TABLA 7.11. *Estudios Retrospectivos* (Adaptación de: Cohen y Manion, 2002).

Características	Fortalezas	Debilidades
ANÁLISIS RETROSPECTIVO DE LOS SUJETOS DE LA MUESTRA	<ul style="list-style-type: none"> • Se utiliza para establecer relaciones causales. • Los datos no han sido modificados ni afectados por agentes externos. 	<ul style="list-style-type: none"> • El recuerdo de la información puede estar incompleto o distorsionado. • Los sujetos interpretan sus recuerdos a partir de los acontecimientos actuales. • Es difícil separar las causas reales y las percepciones personales. • No se puede confirmar ni contrastar.

FUENTES DOCUMENTALES

Como punto final de este capítulo, hemos creído oportuno presentar una serie de páginas web que pueden acercar más al lector a los estudios de encuestas, tanto a su práctica en el ámbito educativo como a sus posibilidades en general.

De entre todas las revistas en formato papel destacamos dos: *Revista de Investigación Educativa (RIE)*, *Revista de Metodología de las Ciencias del Comportamiento*.

En cuanto a revistas electrónicas, destacamos:

- *RELIEVE* Revista electrónica de investigación y evaluación educativa consultable en (<http://www2.uca.es/RELIEVE>). Editada por AIDIPE, Asociación Interuniversitaria de Investigación Pedagógica.
- *Metodología de encuesta* (<http://www.sipic.org/mc>). Revista electrónica de Metodología aplicada a la Psicología y la Educación. La publica la asociación SIPIE (*Sociedad Internacional de Profesionales de la Investigación por Encuestas*).

- REMA (<http://www3.uniovi.es/user.html/herrero/REMA>) Revista electrónica de metodología aplicada a la psicología y la educación, publicada en la Universidad de Oviedo. Permite bajar programas informáticos elaborados por investigadores.
- *Página de Instrumentos de Evaluación. Prometeo* (<http://prometeo.us.es/recursos/>). En esta página se pueden encontrar materiales de acceso gratuito, desde cómo hacer un currículum, hasta diferentes instrumentos de recogida de información, cuestionarios con preguntas abiertas y cerradas, diferenciales semánticos, escalas de valoración, etc.

EJERCICIOS

1. Resume brevemente en qué situaciones es más adecuado utilizar los estudios de encuesta. Pon algunos ejemplos.
2. Suponte que tienes que realizar un estudio sobre la inserción laboral de los graduados/as del ámbito de la educación de tu universidad. ¿Qué tipo de encuesta utilizarías y por qué? Para responder confecciona una tabla donde aparezcan las fortalezas y debilidades de cada uno de los tipos de encuesta adaptados al tema de estudio.
3. Seguro que alguna vez has sido objeto de una encuesta. A partir del contenido que se presenta en este capítulo intenta identificar los aspectos relevantes que recuerdes, como por ejemplo:

Temática de la encuesta

Cómo te seleccionaron para la entrevista

Tipo de encuesta

¿Crees que utilizaron el tipo de encuesta más apropiado para el objetivo de la investigación? ¿Te motivaron lo suficiente para responder sinceramente y desinteresadamente?

Describe brevemente cómo lo hubieses enfocado tú.

CAPÍTULO 8

EL ANÁLISIS CUANTITATIVO DE LOS DATOS

Ruth Vilà Baños
Rafael Bisquerra Alzina

OBJETIVOS

- a) Situar el papel de la estadística en el análisis de datos.
- b) Presentar algunas de las aportaciones de la informática aplicada a la investigación educativa, especialmente haciendo referencia al paquete estadístico del SPSS.
- c) Ofrecer algunos procesos de depuración de la matriz de datos.
- d) Introducir algunos conceptos sobre los valores perdidos en una matriz de datos.
- e) Situar la estadística descriptiva en el análisis exploratorio de los datos.
- f) Presentar algunas valoraciones acerca de la utilización de pruebas paramétricas o no paramétricas.
- g) Clarificar algunos aspectos sobre la significación estadística que ofrece SPSS en las pruebas de contraste.
- h) Presentar alguna de las tendencias actuales en el análisis exploratorio de los datos.
- i) Enumerar algunas aplicaciones de la estadística inferencial y del análisis multivariante en el análisis de los datos.

LA ESTADÍSTICA EN EL ANÁLISIS DE DATOS

Una fase del proceso de investigación científica es el análisis de datos. En este capítulo nos centramos en la metodología cuantitativa, donde la estadística es la técnica de análisis de datos cuantitativa que ha jugado un papel preponderante, por no decir exclusivo. Por este motivo hemos dedicado este capítulo a pre-

sentar una panorámica general acerca del papel que juega la estadística aplicada a la investigación educativa.

El concepto de estadística ha ido variando históricamente, pasando por numerosas concepciones y definiciones. En la actualidad la estadística tiene por objeto el estudio de técnicas y métodos para el tratamiento de conjuntos de datos numéricos, permitiendo la descripción y el análisis inferencial sobre conjuntos de datos, y pudiendo ser aplicada en distintos campos del saber, entre ellos la educación. Se distingue así, entre la estadística formal, teórica, matemática o pura y la estadística aplicada a la investigación y al análisis de la realidad. Para conocer en profundidad los contenidos que aquí se citan sobre análisis de datos, recomendamos la consulta de manuales como los de Etxebarria y Tejedor (2004) sobre *Análisis descriptivo de datos en Educación* y Tejedor y Etxebarria (2004) sobre *Análisis inferencial de datos en Educación*.

FASES EN EL ANÁLISIS DE DATOS

La estadística tiene un papel fundamental en gran parte del proceso de investigación. Como por ejemplo en el planteamiento del problema y formulación de hipótesis que condicionarán en gran medida el tipo de análisis a realizar; en revisiones bibliométricas; en estudios piloto previos a una investigación; en la creación de instrumentos de medida para asegurar su validez y fiabilidad; y sobretodo en lo que respecta al análisis de datos para llegar a conclusiones.

En la etapa del análisis cuantitativo de datos se pueden distinguir tres etapas fundamentales:

- Una fase de análisis exploratorio inicial de los datos que consiste en la depuración de la matriz de datos y en análisis descriptivos de los mismos.
- Una segunda fase de análisis bivariable a través de estudios inferenciales.
- Una tercera fase basada en análisis multivariantes.

En el *análisis exploratorio inicial*, también llamado EDA (exploratory data análisis), tiene especial importancia la *depuración de la matriz* de datos y primera aproximación al análisis descriptivo univariable a través de distribuciones de frecuencia, representacio-

nes gráficas univariadas, medidas de tendencia central, de variabilidad o dispersión, de posición, de asimetría, de curtosis, de comprobación de supuestos paramétricos, etc.

Las pruebas de decisión estadística constituyen un aspecto importante del análisis de datos cuantitativo. Ejemplos de estas pruebas son la correlación de Pearson, prueba t de Student, análisis de la varianza (ANOVA), pruebas no paramétricas como χ^2 (ji-cuadrado), T de Wilcoxon, U de Mann-Whitney, prueba de McNemar, etc.

El análisis multivariante es propio de los estudios más sofisticados. Se justifica en el principio de *causación múltiple*, es decir, que los fenómenos complejos obedecen a múltiples causas y no a una sola. Esto tiene especial importancia, dado que las investigaciones en ciencias sociales y especialmente en el área educativa se ven sometidas a un gran número de variables que intervienen. Algunos ejemplos de análisis multivariante son el análisis factorial de la varianza (ANOVA), análisis de la covarianza (ANCOVA), análisis multivariante de la varianza (MANOVA), el análisis factorial (AF) de componentes principales, el *cluster analysis*, análisis discriminante, etc.

Hasta los años setenta el análisis de datos se hacía prácticamente a mano o con la ayuda de una calculadora. Esto desanimaba la aplicación de análisis multivariados por la complejidad del cálculo. Con la llegada de los ordenadores personales en los años ochenta se pueden abordar análisis que antes eran impensables. Curiosamente esto coincide con el progresivo énfasis en la metodología cualitativa, donde los análisis estadísticos están ausentes o son muy elementales.

LA INFORMÁTICA APLICADA A LA INVESTIGACIÓN EDUCATIVA: EL PAQUETE ESTADÍSTICO SPSS

Actualmente el análisis de datos se realiza por ordenador, mediante paquetes estadísticos. Existen muchos paquetes estadísticos: SAS, BMDP, SPSS, Statgraphics, etc. Uno de los más reconocidos y utilizados en investigación social es el SPSS (Statistical Package for Social Sciences). Los paquetes estadísticos incorporan innovaciones permanentemente. Las figuras e instrucciones que se citan en este capítulo corresponden a la versión 11.0 del SPSS. Esto exige que para profundizar en las aportaciones y el

manejo del SPSS, o de cualquier paquete estadístico, se tenga que recurrir a la consulta de manuales específicos y actualizados. A partir de los años noventa estos manuales se han hecho muy accesibles (por ejemplo: Pardo y Ruiz, 2002).

LA MATRIZ DE DATOS

Una vez finalizado el proceso de recogida de datos, toma protagonismo la elaboración de la matriz de datos (que se encuentra en el visor de edición de datos del paquete SPSS). Una matriz de datos es una estructura en forma de tabla que contiene los valores de cada sujeto en diferentes variables. Habitualmente se colocan las variables en columnas y los individuos en filas. En la fig. 8.1. se puede contemplar un ejemplo de matriz de datos.

Para elaborar una matriz de datos hay que tener muy claro el concepto de variable. Por ejemplo, si en la matriz de datos se analiza un test de 50 ítems, cada ítem funciona como una variable. Además hay una variable que es la puntuación total. Es posible que en el análisis solo interese esta última; pero las otras permiten un análisis de ítems que es recomendable.

La matriz de datos se puede elaborar de diferentes formas: a) entrando los datos directamente a través del teclado; b) mediante lectora óptica; c) utilizar una matriz de datos previa y por tanto entrarla a través de procesos informáticos (disquete, CD-ROM, internet). En el primer caso se pueden producir errores que conviene prevenir y corregir. Por esto, el primer paso del análisis debe ser la *depuración de la matriz de datos*. La consistencia y el valor del análisis de datos se fundamenta en una matriz que a ser posible, no tenga errores. Errores propios que se cometen al escribir en un teclado; por ejemplo, el valor que hay que introducir es 18, y por error se escribe 81; o el uno no se pulsa bien y no queda escrito, con lo que se ha consignado solo 8.

	sexo	región	feliz	vida	herm	hijos	edad	educ	educpad	educma	educasp	presc
1	Mujer	Nor-Este	Muy feliz	Excitante	1	2	61	12	No pr	12	No pro	2
2	Mujer	Nor-Este	Bastante fel	Excitante	2	1	32	20	20	18	20	7
3	Hombre	Nor-Este	Muy feliz	No proced	2	1	35	20	16	14	17	5
4	Mujer	Nor-Este	No contesta	Rutinaria	2	0	26	20	20	20	No pro	4
5	Mujer	Nor-Este	Bastante fel	Excitante	4	0	25	12	No sa	No sa	No pro	4
6	Hombre	Nor-Este	Bastante fel	No proced	7	5	59	10	8	6	No pro	NP, NC
7	Hombre	Nor-Este	Muy feliz	Excitante	7	3	46	10	8	No sa	No pro	NP, NC
8	Mujer	Nor-Este	Bastante fel	No proced	7	4	No conte	16	5	6	No pro	E
9	Mujer	Nor-Este	Bastante fel	Rutinaria	7	3	57	10	6	5	No pro	NP, NC
10	Mujer	Nor-Este	Bastante fel	Excitante	1	2	64	14	8	12	20	3
11	Hombre	Nor-Este	Bastante fel	Excitante	6	0	72	9	12	No sa	No pro	3
12	Mujer	Nor-Este	Muy feliz	No proced	2	5	67	12	8	8	13	2
13	Hombre	Nor-Este	Bastante fel	No proced	1	0	33	15	11	12	14	E
14	Hombre	Nor-Este	Bastante fel	Rutinaria	2	1	23	14	12	12	No pro	4
15	Mujer	Nor-Este	Bastante fel	Rutinaria	7	1	33	12	12	12	No pro	E
16	Mujer	Nor-Este	Muy feliz	Rutinaria	6	2	59	12	8	No sa	12	NP, NC
17	Hombre	Nor-Este	Bastante fel	No proced	4	1	60	14	6	6	No pro	3
18	Hombre	Nor-Este	Muy feliz	Rutinaria	6	2	77	9	0	0	8	3
19	Mujer	Nor-Este	Bastante fel	No proced	12	2	52	14	8	12	8	E
20	Hombre	Nor-Este	Muy feliz	Aburrida	5	1	55	7	No sa	No sa	16	4
21	Mujer	Nor-Este	Muy feliz	Rutinaria	2	1	37	14	12	12	No pro	4
22	Mujer	Nor-Este	Bastante fel	No proced	7	0	45	9	8	No co	No pro	NP, NC

Fig. 8.1. Ejemplo de Matriz de datos con el paquete estadístico SPSS.

La sugerencia de depurar la matriz de datos consiste básicamente en calcular algunos descriptivos de todas las variables, como por ejemplo:

- *El mínimo y el máximo* para verificar que no existan valores extremos que sobrepasen o no alcancen el rango lógico de la variable. Por ejemplo, si tenemos la variable edad de los estudiantes universitarios se trataría de verificar que no exista ningún valor inferior a 18; valores superiores a 50 podrían ser posibles, aunque sospechosos. Si se observa un valor de 81, aunque es teóricamente posible, convendría comprobar que no se trata en realidad de un error de inversión: se ha escrito 81 en lugar de 18. En este caso hay que corregirlo antes de pasar al análisis de datos.
- *El número de individuos* de cada variable para verificar que sea correcto y no se ha olvidado o duplicado a ningún individuo.

En la depuración de la matriz de datos se trata de asegurar que los análisis posteriores se van a basar en unos datos reales. Al

no detectarse los errores de transcripción, podrían producirse errores graves en los resultados.

Por otra parte, algunos de estos procesos de depuración de la matriz se inician de forma previa a la introducción de datos: cuando se definen las variables que van a formar la matriz. En el SPSS existe el proceso de *definir variables* donde se permite especificar aspectos de la variable como su nominación, la escala de medida, el tipo de variable, las categorías que obtiene en caso de ser categórica, aspectos de formato en su presentación (anchura, alineación, decimales), así como los valores perdidos, tal como se presenta en la figura 8.2. Los *valores perdidos* son aquellos casos en que no se dispone de información sobre el valor de una variable determinada para algún individuo.

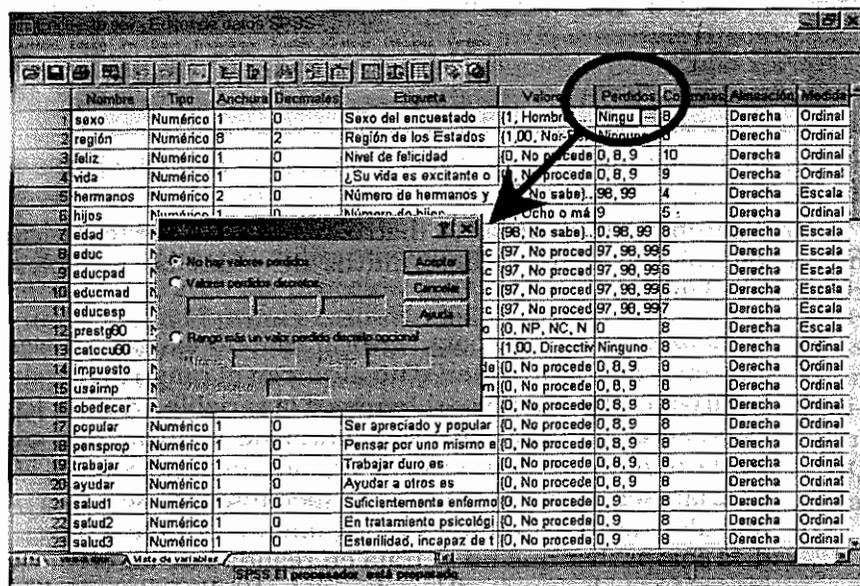


Fig. 8.2. Ejemplo de pantalla de vista de variables con la opción de valores perdidos activada.

ANÁLISIS EXPLORATORIO DE LOS DATOS

Con posterioridad a la depuración de la matriz de datos suele priorizarse un análisis exploratorio (EDA¹) basado en el análisis descriptivo de cada variable. Esto permite una mayor comprensión del fenómeno.

Tiene especial importancia la detección de *outliers* o sujetos cuyas puntuaciones difieren marcadamente de los modelos establecidos por el resto de individuos de la muestra.

La estadística descriptiva univariable, así como la comprobación de los supuestos paramétricos, completan el análisis exploratorio de los datos. En la figura 8.3. se presenta un ejemplo de output del SPSS de un análisis descriptivo de una variable.

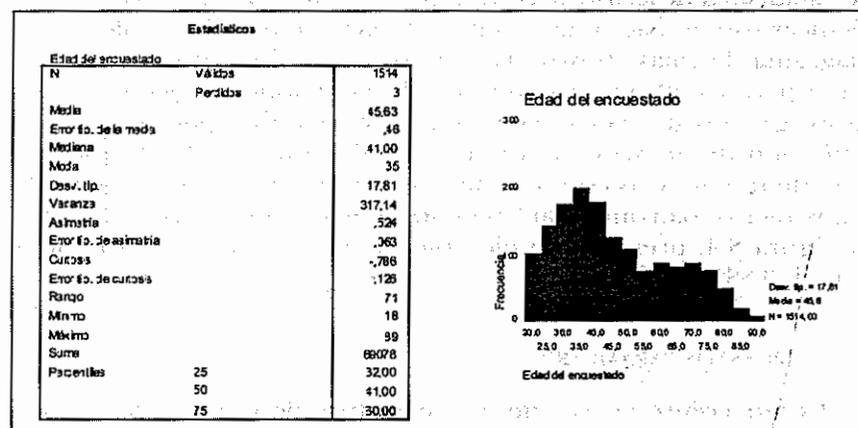


Fig. 8.3. Ejemplo de output del SPSS sobre análisis descriptivo de una variable.

Este ejemplo responde al comando *Frecuencias*² del programa para los estadísticos descriptivos que se presentan:

- Indicadores de tendencia central: media aritmética, error típico de la media, moda, mediana.

¹ Exploratory Data Analysis.

² El comando *Frecuencias* se obtiene desde la opción *Analizar* del menú principal donde se ejecutan todos los análisis estadísticos, en la opción de *Estadísticos descriptivos*.

- Indicadores de dispersión: varianza, desviación típica, rango, mínimo, máximo.
- Indicadores de posición: cuartiles, deciles, percentiles³.
- Indicadores de forma: asimetría, y su error típico, curtosis, y su error típico.
- Representación gráfica: histograma con la hipotética curva normal superpuesta.

Para profundizar en estos elementos, se recomienda consultar manuales de estadística básicos que expliquen con mayor profundidad aspectos de estadística descriptiva.

La expresión EDA (Exploratory Data Analysis) se puso en circulación en los años ochenta con la propuesta de nuevas técnicas de estadística descriptiva asistida por ordenador. Algunas de estas técnicas son el diagrama de tallo y hojas (stem-and-leaf plot), y el diagrama de cajas (box-plot). Ambos procedimientos son de utilidad para examinar visualmente distribuciones de valores para varios grupos de datos. Estas técnicas pueden ser útiles para la detección de errores en la matriz de datos, observar la distribución de frecuencias con más detalle (valores extremos, rangos de datos vacíos, patrones, variabilidad de los datos inesperada), etc. La figura 8.4. presenta un ejemplo de ambas gráficas elaboradas con el SPSS⁴.

LOS SUPUESTOS PARAMÉTRICOS

Como continuación de la estadística descriptiva para cada variable suele interesar la comprobación de los supuestos paramétricos para determinar si es procedente aplicar pruebas paramétricas o si son preferibles las no paramétricas para los datos disponibles. Cada técnica de análisis estadístico requiere el cumplimiento de unas condiciones. El no cumplimiento de estas condiciones puede suponer errores en el análisis de datos.

Los *supuestos paramétricos* que conviene confirmar son:

- Muestras grandes: en cada grupo $n > 30$.

³ El output siempre muestra percentiles, aunque tiene la opción de pedir cualquiera de los indicadores decil o cuartil.

⁴ Mediante el procedimiento *Analizar* del menú, en su opción *estadísticos descriptivos* y dentro de ésta, la opción *explorar*.

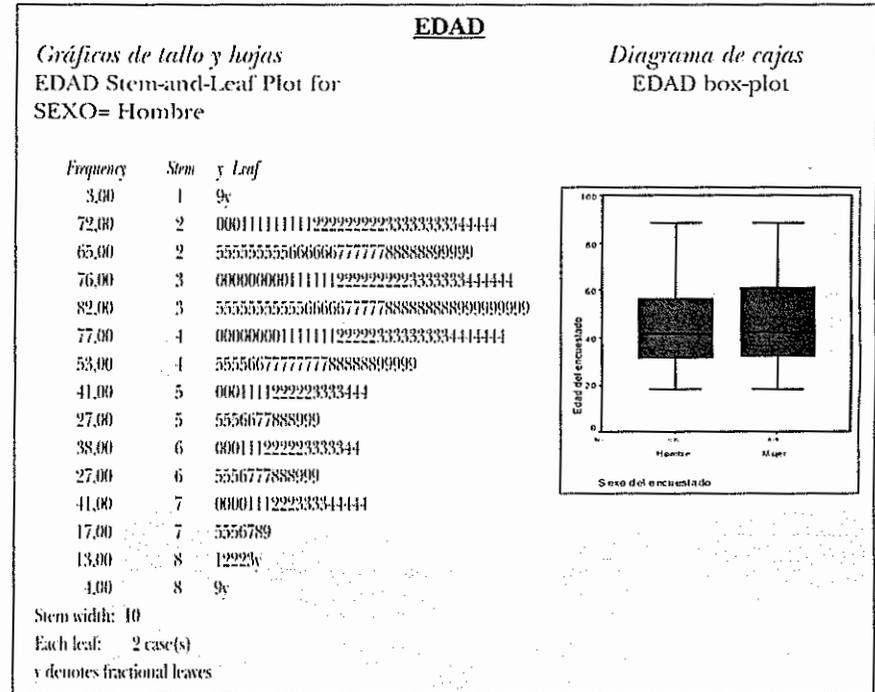


Fig. 8.5. Ejemplo de diagrama de tallo y hojas y diagrama de cajas.

- Normalidad: al aplicar la prueba de Kolmogorov-Smirnov (K-S) se confirma que la variable se ajusta a la distribución normal.
- Homocedasticidad: hay homogeneidad de varianzas en grupos independientes (prueba F de Fisher, de Levene, de Box, etc.).
- Linealidad: la correlación entre pares de variables es lineal.

Si se confirman estos supuestos paramétricos se pueden aplicar pruebas paramétricas, que suelen tener más ventajas respecto a las no paramétricas.

Se ha observado que en la práctica muchos supuestos paramétricos no se cumplen *strictu sensu*, pero “casi” se cumplen. Por esto, aunque no se confirmen estrictamente todos los supuestos paramétricos, se pueden aplicar pruebas paramétricas, conocien-

do los efectos que ello pueda suponer sobre los resultados. Desde diversas perspectivas, se aconseja el uso de pruebas paramétricas siempre que sea posible, ya que hay razones teóricas y empíricas que justifican que muchas pruebas paramétricas son lo suficientemente robustas para que los resultados no se vean seriamente alterados por ligeras desviaciones de los supuestos paramétricos.

Las pruebas no paramétricas aportan una solución para las situaciones en que no se cumplen ostensiblemente los supuestos paramétricos. Para profundizar en las pruebas no paramétricas recomendamos la consulta de tratados específicos.

En la figura 8.5 se pretende verificar uno de los supuestos paramétricos: el de la normalidad. Para ello se ha aplicado la prueba de Kolmogorov-Smirnov para una muestra⁵. Esta prueba, al igual que el resto de contrastes se interpreta a través del grado de significación *p*, tal como se visualiza en el output: sig. asintót. (bilateral). Es decir, "significación asintótica bilateral", que significa la probabilidad de error al rechazar la hipótesis nula en una prueba bilateral. Antes de avanzar más en el análisis exploratorio, es necesario introducir unas notas sobre la significación estadística.

Prueba de Kolmogorov-Smirnov para una muestra		
		Edad del encuestado
N		1514
Parámetros nomales ^{a,b}	Media	45,63
	Desviación típica	17,81
Diferencias más extremas	Absoluta	,105
	Positiva	,105
	Negativa	-,067
Z de Kolmogorov-Smirnov		4,090
Sig. asintót. (bilateral)		,000

a. La distribución de contraste es la Normal.
b. Se han calculado a partir de los datos.

Fig. 8.5. Prueba de Kolmogorov - Smirnov.

⁵ La prueba de Kolmogorov - Smirnov se encuentra en *Analizar* del menú principal, en la opción de *pruebas no paramétricas*, bajo la denominación de *K - S de 1 muestra*.

SIGNIFICACIÓN ESTADÍSTICA

Al aplicar cualquier prueba de decisión estadística se establece previamente un *nivel de significación*, que representa la *probabilidad de error* que se está dispuesto a asumir al rechazar la hipótesis nula. En Ciencias Sociales los niveles de significación más usuales son el 0'05 y el 0'01.

Estos niveles tradicionalmente eran consultados en las respectivas tablas estadísticas (según la prueba efectuada) que daba paso a la interpretación de los resultados. No obstante, en la actualidad, en los outputs proporcionados por los paquetes estadísticos como el SPSS, se proporciona el grado de significación *p* calculado a posteriori (Bisquerra, 1987: 92-93). Mediante la *p* no es necesaria la consulta de tabla alguna, ya que su interpretación es inmediata: se compara la *p* del *output* con el nivel de significación elegido a priori (0'05 o 0'01). Por ejemplo: en la figura 8.4 para la prueba de Kolmogorov - Smirnov se obtiene una "sig. asintót. (bilateral)" de 0.000. Al comparar este valor con el 0'05 (elegido previamente), se observa que $0.000 < 0'05$. Esto permite concluir que se rechaza la hipótesis nula, ya que la probabilidad de error al hacerlo es inferior al riesgo que estábamos dispuestos a asumir. En este caso, rechazar la hipótesis nula significa rechazar la hipótesis de normalidad. Es decir, la distribución no se ajusta a la distribución normal.

Hay que tener en consideración que a menudo es fácil conseguir significación estadística (rechazar la hipótesis nula) en muestras muy grandes (con miles de sujetos). Por esto conviene distinguir entre significación estadística y significación sustantiva. Por ejemplo, encontrar una diferencia estadísticamente significativa entre la media de puntuaciones de un grupo de 5.5 y otro grupo de 5.7, quizás no sea educativamente significativo. De este modo, se recomienda que el análisis de significación debe hacerse a un doble nivel: significación estadística y significación sustantiva. En este último caso, se trata de un análisis cualitativo complementario que vaya más allá de los números. Los números son indicadores que muestran tendencias; las aportaciones de la metodología cualitativa a veces tienen aplicación en el análisis cuantitativo de datos, para llegar a una interpretación de los resultados que no se limite al dato numérico, sino que tome en consideración aspectos cualitativos como la magnitud del fenómeno, la significación sustantiva, el

marco teórico, aplicaciones prácticas, la realidad del fenómeno, etc. La integración y complementariedad de métodos cuantitativos y cualitativos puede llegar al análisis de datos cuantitativos, superando su tradicional rigidez.

LA ESTADÍSTICA INFERENCIAL Y EL ANÁLISIS MULTIVARIANTE

Un paso en el análisis de datos puede ser el estudio entre pares de variables (bivariable), como la correlación o las pruebas de contraste de medias, tanto paramétricas (t de Student, ANOVA) como no paramétricas (ji-cuadrado, t de Wilcoxon, etc.). Para profundizar en el análisis descriptivo de datos, así como en el inferencial, se recomienda consultar manuales como los de Etxeberria y Tejedor (2004) y Tejedor y Etxeberria (2004).

La interpretación de las pruebas de significación estadística, tal como se ha comentado en el apartado anterior, se lleva a término mediante el grado de significación p , que representa la probabilidad de error que se está dispuesto a asumir al rechazar la hipótesis nula. En este sentido, el conocimiento al que se llega a través de la estadística es probabilístico. Siempre se contempla un margen de error. Esta es una de las características de la estadística, ya que el conocimiento que proporciona no es absoluto. Es importante ser conscientes de ello en el análisis de resultados y establecimiento de conclusiones, para evitar un uso incorrecto de la estadística.

La complejidad de los fenómenos educativos pone de manifiesto limitaciones al utilizar pocas variables y aplicar únicamente pruebas bivariadas. En educación intervienen un gran número de variables, que pueden ser analizadas de forma más compleja y global con una aproximación multivariante. El avance que supone la informática aplicada a la investigación educativa permite una mayor complejidad en los procesos de análisis estadístico. En este sentido, el análisis multivariante se reafirma como un conjunto de técnicas que, a pesar de su complejidad matemática, pueden ser fácilmente aplicadas a través de paquetes estadísticos, como el SPSS. De forma simplificada, podemos afirmar que el análisis multivariante es aquel que se aplica a más de dos variables simultáneamente. Profundizar en el análisis multivariante excede el marco de este capítulo; por esto remitimos a manuales específicos sobre el tema, como por ejemplo los de Bisquerra (1989) y Visauta (1998).

EJERCICIOS

1. ¿Cuáles crees que son las aportaciones de la informática aplicada a la investigación educativa?
2. ¿En qué consiste la depuración de la matriz de datos? ¿Qué sentido tiene realizar la depuración de la matriz de datos?
3. ¿Qué importancia tienen los valores perdidos en el proceso de depuración de la matriz de datos?
4. ¿En qué consisten las tendencias actuales en el análisis exploratorio de los datos?
5. ¿Qué supone la estadística descriptiva en el análisis de datos cuantitativo?
6. ¿Qué sentido tiene verificar los supuestos paramétricos en los datos?
7. ¿Cómo se lleva a cabo la interpretación de la significación p ?
8. ¿Qué supone la estadística inferencial en el análisis de datos cuantitativo?
9. ¿Cuáles crees que son las aportaciones de la informática aplicada a la investigación educativa en el análisis multivariante?
10. En una matriz de datos propia realiza el proceso de análisis de datos siguiente. (Este ejercicio implica una clara implicación de la estadística que va más allá de las informaciones aportadas por este manual; requiere una interrelación con la clase de Estadística):
 - Depuración de la matriz de datos.
 - Análisis exploratorio de los datos.
 - Comprobar el supuesto paramétrico de normalidad.
 - Realiza un contraste de medias.