

En este diseño se aplican todos los tratamientos a todos los grupos que intervienen en la investigación. Generalmente este tipo de diseños se utilizan cuando hay varios tipos de tratamientos que se desea analizar. La novedad está en que cada grupo es expuesto al tratamiento en momentos distintos tal y como se puede apreciar en el diagrama. En éste observamos que el número de tratamientos ha sido de tres y el orden en que los grupos han sido expuestos varía.

En el caso de que los sujetos pudiesen asignarse al azar a los grupos, estaríamos ante un diseño de corte experimental. El uso de este tipo de diseño es apropiado en aquellos casos en los que la aplicación de un tratamiento no afecta a la aplicación del siguiente tratamiento. El estudio de los efectos que la ingesta de determinadas sustancias puede tener en la concentración o atención de los sujetos al llevar a cabo una tarea es un claro ejemplo de aplicación de este tipo de diseños. Si un investigador quisiese estudiar el efecto que la cafeína produce en la capacidad de concentración de los alumnos, en la realización de tareas como el tachado de letras como una prueba de atención, el tratamiento podría aplicarse a los mismos grupos suministrando cantidades distintas de cafeína en cada uno de los tratamientos. El investigador podría incluso usar un placebo como uno de los tratamientos.

La secuencia a seguir en el desarrollo de este diseño sería: aplicación de todos los tratamientos a todos los grupos en orden diferente dejando un espacio de tiempo, medida de la variable dependiente después de cada tratamiento. Una vez finalizado el proceso, se procede a analizar los resultados:

Réplicas	Tratamientos		
	X_1	X_2	X_3
1	a0	b0	c0
2	b0	c0	a0
3	c0	a0	b0
Media	$\bar{X}_0$	$\bar{X}_0$	$\bar{X}_0$
X_1 = Placebo X_2 = Un café X_3 = Dos cafés 0 = Prueba de atención Grupos a, b y c Réplicas 1, 2 y 3			

En la réplica 1 el tratamiento X_1 es aplicado al grupo a, el tratamiento X_2 al grupo b y el tratamiento X_3 al grupo c. En la réplica 2 el tratamiento X_1 es aplicado al grupo b, el tratamiento X_2 al grupo c y el tratamiento X_3 al grupo a. En la réplica 3 el tratamiento X_1 es aplicado al grupo c, el tratamiento X_2 al grupo a y el tratamiento X_3 al grupo b. A continuación se calcula la media de la actuación de los sujetos en cada uno de los tres tratamientos (media de columnas).

El análisis estadístico de los resultados de este tipo de diseños se pueden llevar a cabo a través del análisis de la varianza.

6. LOS DISEÑOS FACTORIALES

Anteriormente hemos descrito que el diseño clásico de investigación mantiene todas las variables constantes excepto la variable independiente que es aplicada al grupo experimental. Cuando el investigador finaliza su investigación, si la media de la variable dependiente es distinta en el grupo experimental frente al grupo de control, puede concluir, en términos generales, que el tratamiento ha producido los efectos más o menos esperados. Sin embargo, análisis estadísticos más detallados podrían estar indicando que las ganancias observadas en la variable dependiente, en el grupo experimental, en el subgrupo de mujeres es mayor que en el subgrupo de hombres. Esta observación puede llevar a pensar que el tratamiento no es independiente de la variable género, sino que entre ambas hay un efecto de interacción.

Recordarás que en el Capítulo 2, cuando describíamos las variables, mencionamos tres tipos de variables independientes: la que hace de tratamiento o causa, la moderador y la variable control. En el ejemplo que acabamos de mencionar, la variable género desempeñaría la función de variable moderador puesto que el investigador la incluye en sus estudios para ver cuáles son sus efectos en combinación con el tratamiento.

Las técnicas estadísticas que hemos mencionado hasta aquí son de bastante utilidad para algunos de los diseños ya expuestos, como por ejemplo la t de Student o el análisis de varianza simple o factorial. Sin embargo, cuando el investigador incluye en su estudio más de una variable independiente es necesario tratar los datos con otro tipo de técnicas que permitan analizar no sólo a la variable independiente principal o tratamiento, sino las otras variables independientes de carácter moderador o diferenciador. Los diseños factoriales son los que posibilitan el estudio conjunto de más de una variable independiente sobre la variable dependiente. El análisis de varianza factorial analizará el efecto de estas variables por separado (efectos principales) y de forma conjunta (efectos de interacción).

En las páginas que siguen vamos a presentar algunos de los diseños factoriales que el investigador puede utilizar dependiendo del número y tipo de variables moderador que incluya en su estudio. Junto a lo que aquí vamos a exponer recomendamos al lector que haga un repaso del análisis de la varianza factorial en cualquier libro de estadística (Weilkowitz, Ewen y Cohen, 1981). La explicación de los ANOVA factoriales le serán de una gran ayuda para proceder en este tipo de investigaciones. Aquí sólo abordaremos dos tipos de ANOVA, el de dos factores y el de tres factores.

6.1. Diseño de ANOVA dos factores

En este diseño una de las variables independientes es considerada como el tratamiento, y la otra, la variable moderador, que podría ser otro tratamiento o una característica de los sujetos o del ambiente. El ANOVA dos factores permite analizar cada una de tales variables por separado (efectos principales) y de forma conjunta (efecto de interacción). Este diseño exige que los sujetos hayan sido asignados al azar a cada uno de los niveles del tratamiento.

Los dos factores o variables independientes pueden ser de carácter cualitativo o cuantitativo o una combinación de ambos. Por ejemplo, supongamos que un investigador desea conocer el efecto que la caféina tiene sobre la capacidad de atención que los alumnos de bachillerato necesitan en tareas escolares. Con este fin desarrolla un estudio en el que a cuatro grupos de sujetos aplica un nivel distinto de tratamiento. El factor tratamiento es la cantidad de caféina (variable cuantitativa) suministrada a cada uno de los subgrupos. Dado que el investigador tiene la sospecha de que el género (variable cualitativa) de los sujetos puede ser una variable relevante, la incluye como variable moderador. Si los niveles del tratamiento son cuatro (placebo, un descafeinado, un café normal, dos cafés) y el género (hombres y mujeres), estamos ante un diseño factorial 4×2 . El investigador analizará los efectos principales de cada una de las dos variables independientes y el efecto conjunto de ambas variables combinando cada nivel en cada uno de los factores. La variable dependiente podría ser una prueba de atención después de ser sometidos los sujetos al tratamiento (Wynne, 1982).

La anatomía del diseño queda reflejada en el gráfico siguiente:

		A			
		A1	A2	A3	A4
B	B1	A1B1	A2B1	A3B1	A4B1
	B2	A1B2	A2B2	A3B2	A4B2

A = Cantidad de caféina
 A1 = Placebo
 A2 = Un café descafeinado
 A3 = Un café normal
 A4 = Dos cafés

B = Género
 B1 = Mujeres
 B2 = Hombres

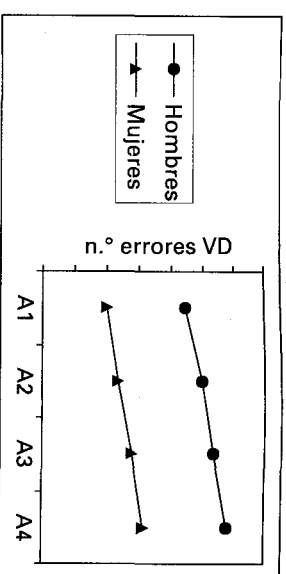
En este estudio el investigador evaluará tres tipos de efectos:

- La diferencia que existe en los resultados de la prueba de atención (VD) después del tratamiento entre los cuatro grupos en función del tratamiento (VI) al que han sido sometidos.
- Diferencia en la prueba de atención (VD) entre los hombre y la mujeres (VI).
- Diferencia entre los ocho grupos en función de la combinación de cada uno de los niveles de los dos factores. Es decir, ¿el resultado en la prueba de atención depende del género y el tratamiento, de forma conjunta, al que han sido sometidos los participantes?

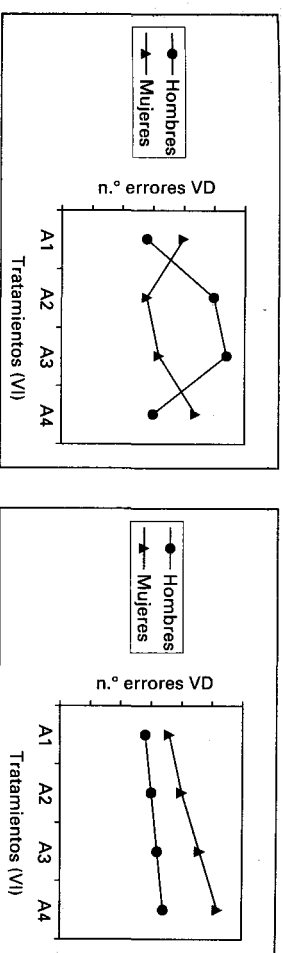
Como podemos apreciar, el investigador estudia por un lado los efectos principales, la diferencia en función del tratamiento (A1, A2, A3 y A4) y la diferencia en función del género (B1 y B2). Pero lo realmente importante en este tipo de diseños es el estudio de los efectos de interacción (A1B1, A1B2, etc.).

Estos efectos de interacción deberán ser representados gráficamente y pueden adoptar las formas siguientes:

No interacción:



Interacción:



Ejemplo con datos para un ANOVA de dos factores: 4×3 .

Fuente	Suma de cuadrados	Grados de libertad	Cuadrado medio	F
A	60,16	$a - 1 = 3$	20,05	2,94*
B	78,78	$b - 1 = 2$	39,39	5,76*
AB	18,88	$(a - 1)(b - 1) = 6$	3,15	0,46
Error	409,17	$ab(n - 1) = 60$	6,82	
Total	566,99	$abn - 1 = 71$		

* Significación estadística al 0,05.

Los datos de la tabla anterior expresan los resultados de un estudio (Collyer y Enns, 1991). Estos datos corresponden a un diseño 4×3 . En la tabla podemos apreciar dónde se producen las diferencias significativas: en los efectos principales, es decir, entre los niveles del factor A (entre A1, A2, A3 y A4); entre los niveles del factor B (B1, B2 y B3); y, por último, en las interacciones de los niveles de los dos factores. El asterisco expresa el nivel de significación. Para el ejemplo que presentamos sólo aparecen significativos los efectos principales pero no los de interacción.

6.2. Diseño de ANOVA tres factores

Siguiendo con el mismo ejemplo (Collyer y Enns, 1991) que acabamos de ver, supon- gamos que el investigador incluye un tercer factor o variable independiente como puede ser el momento en el que el investigador suministra el tratamiento a los sujetos, por la mañana o por la tarde. En este caso tendríamos un diseño de $4 \times 2 \times 2$.

La anatomía del diseño sería la que sigue:

A									
		A1		A2		A3		A4	
		C1	C2	C1	C2	C1	C2	C1	C2
B	B1	A1B1C1	A1B1C2	A2B1C1	A2B1C2	A3B1C1	A3B1C2	A4B1C1	A4B1C2
	B2	A1B2C1	A1B2C2	A2B2C1	A2B2C2	A3B2C1	A3B2C2	A4B2C1	A4B2C2

El investigador evaluará los siguientes tipos de efectos:

- a) La diferencia que existe en los resultados de la prueba de atención (VD) realizada después del tratamiento entre los cuatro grupos en función del tratamiento al que han sido sometidos (A1, A2, A3 y A4).
- b) Diferencia entre los dos niveles del segundo factor, el género (B1, B2 y B3).
- c) Diferencia entre los resultados obtenidos en el tercer factor (C1 y C2) en la prueba de atención según la hora del tratamiento.
- d) Diferencia entre los ocho grupos en función de la combinación de cada uno de los niveles de los dos factores primeros, A y B (cuatro tratamientos por dos géneros). Es decir, ¿el resultado en la prueba de atención depende del género y el tratamiento, de forma conjunta, al que han sido sometidos los sujetos?
- e) Diferencia entre los ocho grupos en función de la combinación de cada uno de los niveles de los otros dos factores, A y C (cuatro tratamientos por hora). Es decir, ¿el resultado en la prueba de atención depende del tratamiento y de la hora en que se aplicó dicho tratamiento, de forma conjunta, al que ha sido sometido el sujeto?
- f) Diferencia entre los cuatro grupos en función de la combinación de cada uno de los niveles de los dos factores, B y C (género por hora). Es decir, ¿el resultado en la prueba de atención depende de la hora y el género, de forma conjunta, al que ha sido sometido el sujeto?
- g) Diferencia entre los 16 grupos en función de la combinación conjunta de los niveles del tratamiento (A), género (B) y hora (C).

Como podemos apreciar, el investigador estudia por un lado los efectos principales, la diferencia en función del tratamiento y la hora en que se aplicó el tratamiento y la diferencia en función del género. Pero lo realmente importante en este tipo de diseños es el estudio de los efectos de interacción.

Queremos llamar la atención del efecto multiplicador que ha supuesto en el número de grupos que se necesitan en los diseños factoriales cuando se introduce un factor más y el número de los niveles del factor es grande. En el diseño factorial de 4×2 el número de grupos necesarios es de 8; al introducir un factor más, con dos niveles, el número de subgrupos se ha duplicado. Este punto es importante resaltarlo porque con la inclusión de más de tres factores, especialmente cuando el número de niveles en cada factor es grande —más de dos—, los diseños factoriales se vuelven muy complejos y el manejo y la interpretación de los resultados especialmente difícil.

Veamos un ejemplo numérico de ambos ANOVAS factoriales (Collyer y Enns, 1991).

ANOVA tres factores:

Fuente	Suma de cuadrados	Grados de libertad	Cuadrado medio	F
A	60,16	$a - 1 = 3$	20,05	2,94*
B	78,78	$b - 1 = 2$	39,39	5,76*
C	136,13	$c - 1 = 1$	136,13	24,80*
AB	18,88	$(a - 1)(b - 1) = 6$	3,15	0,57
AC	1,14	$(a - 1)(c - 1) = 3$	0,38	0,07
BC	4,33	$(b - 1)(c - 1) = 2$	2,17	0,40
ABC	4,24	$(a - 1)(b - 1)(c - 1) = 6$	0,71	0,13
Error	263,33	$abc(n - 1) = 48$	5,49	
Total	566,99	$abcn - 1 = 71$		

* Significación estadística al 0,05.

Los datos de la tabla anterior expresan: los efectos principales, los efectos de interacción de los factores dos a dos y los efectos de interacción de los tres factores a la vez. El asterisco indica dónde aparecen las diferencias significativas. En este ejemplo sólo se han producido datos significativos en los efectos principales; es decir, entre los niveles de los factores A (entre A1, A2, A3, A4); B (entre B1, B2, B3) y C (C1 y C2); pero en ninguna de las interacciones.

7. TIPOS DE VARIABLES TRATAMIENTO EN LOS DISEÑOS EXPERIMENTALES

Las variables independientes que el investigador puede incluir en su investigación, como ya hemos visto anteriormente, pueden clasificarse en variables tratamiento, variables moderador y variables control. Cuando hablamos de la variable independiente no implica que siempre sea «manipulable» por parte del investigador en el sentido de introducir las variaciones que considere oportunas para estudiar cuáles son sus efectos

en la variable dependiente. Algunas veces lo único que puede hacer el investigador es seleccionar determinados valores ya existentes de una variable. Campbell y Stanley han hecho una clasificación de los tipos de variables independientes que podrían ser incluidas en un experimento según su grado de manipulabilidad:

- **Variables manipulables:** método de enseñanza, asignable a los grupos a voluntad del investigador.
- **Variables potencialmente manipulables:** materias que los sujetos cursan y que pueden no estar incluidas en el currículum pero que el investigador las incluye en su estudio.
- **Variables relativamente fijas:** procedentes del ambiente, como por ejemplo los aspectos socioeconómicos.
- **Características orgánicas de los sujetos:** como por ejemplo la edad, el género, la altura, etc.
- **Respuestas de los sujetos a diversas pruebas:** como por ejemplo las puntuaciones de los tests.

Por lo general, en la investigación experimental el investigador está fundamentalmente interesado en las variables del primer apartado. Las de los tres últimos apartados recogen variables que generalmente el investigador las utiliza para generalizar los efectos de la manipulación o el tratamiento. Por ejemplo, al experimentar un nuevo método de enseñanza no sólo desea concluir que dicho método es mejor que otro u otros, sino que dicho método tiene mejores efectos entre los de un género frente a los de otro; o para alumnos con una determinada capacidad intelectual frente a otra, etc.

8. LA INVESTIGACIÓN ATI (INTERACCIÓN ACTITUD-TRATAMIENTO)

Este tipo de investigación se conoce así por las siglas de las palabras inglesas *Aptitude-Treatment Interaction*. Cuando un investigador somete en su estudio a los sujetos a un tratamiento no debe olvidar que los sujetos tienen diferentes estilos de aprendizaje y aptitudes. En las últimas décadas se ha desarrollado una línea de investigación que ha tratado de explorar el emparejamiento entre las características de los estudiantes y los distintos métodos de enseñanza.

La investigación ATI no asume que existan unos métodos de enseñanza mejor que otros, ni tampoco que unos estudiantes, por el hecho de tener unas características determinadas, sean mejores que otros. En su lugar, lo que hace la investigación ATI es establecer que existe una interacción entre métodos y características de los sujetos y que esta interacción puede resultar significativa en algunos niveles de la interacción. Estos efectos de interacción significativa se pueden poner de manifiesto a través del uso de diseños factoriales.

En los diseños factoriales se incluyen normalmente dos tipos de variables. Por un lado, la variable independiente principal, que es el tratamiento o la variable experimental. Esta variable consiste en la aplicación de métodos de enseñanza, materiales de estudio, tipo de escuela, materiales instruccionales, etc. En general, las varia-

bles que el investigador utiliza para este tipo de estudios suelen ser del grupo de las variables susceptibles de manipulación, potencialmente manipulables o variables ambientales relativamente fijas. El segundo tipo de variables independientes que el investigador puede seleccionar se corresponden con características de los sujetos, como, por ejemplo, las características intelectuales, de personalidad, niveles académicos, estilos o enfoques de aprendizaje, etc. Pero no necesaria y exclusivamente las aptitudes que es de donde viene el nombre a esta línea de investigación. Hoy en día los modelos ATI de investigación incluyen en la taxonomía de variables o características de los sujetos no sólo las aptitudes sino cualquier otra característica del aprendiz (Campbell y Stanley, 1973).

9. LOS DISEÑOS DE CASO ÚNICO, $N = 1$

Como su nombre indica, este tipo de diseños se caracterizan porque la muestra de sujetos es una solamente. Esta muestra puede tener desde muy pocos casos (3 o 4) hasta varias decenas. Este tipo de diseños son muy apropiados en investigaciones sobre modificación de conducta.

Los diseños de caso único no deben confundirse con el estudio de casos. Los diseños de caso único se enfrentan con muchas de las características de los diseños vistos hasta aquí, en cuanto al control de las variables, manipulación de la variable independiente, fiabilidad de las observaciones, descripción del tratamiento, control de la validez interna y externa, etc. El estudio de casos, normalmente, se limita a una descripción por parte del investigador que muchas veces puede ser subjetiva porque el objetivo no es tanto la investigación sino la solución de un problema. Muchas investigaciones que han utilizado este tipo de diseños han sido tan rigurosas como cualquiera de los otros diseños descritos.

Dado que en este tipo de diseños no se pueden llevar a cabo algunos de los planteamientos formulados para los diseños anteriores, el investigador puede hacer que su diseño gane en validez interna y externa siguiendo una serie de sugerencias como las que señalamos a continuación (Borg y Gall, 1979).

9.1. Fiabilidad de las observaciones

En este tipo de diseños la observación es una de las técnicas más utilizadas para la recogida de los datos. La primera recomendación es que el investigador siga las normas que la observación exige para que los datos que se recojan sean lo más fiables posible. En el Capítulo 5 de este libro se dedica un amplio espacio a hablar de la observación.

- Medidas repetidas.** Este tipo de diseños exigen medidas repetidas antes del tratamiento y después del tratamiento, con el fin de determinar si los cambios que se producen se deben a la variable independiente y por cuanto tiempo dura la intervención. El objetivo por parte del investigador es que las medidas repetidas de la variable dependiente antes y después del tratamiento implique

a los mismos observadores, se sigan las mismas instrucciones, se mantengan las mismas condiciones ambientales, etc. De otro modo, los efectos del tratamiento se verán contaminados por los efectos de la medida en sí.

b) Descripción de las condiciones experimentales. La descripción del procedimiento seguido por el investigador en su trabajo es condición *sine qua non* para la réplica de las investigaciones y, por consiguiente, para ganar en validez externa. Este tipo de diseños exigen que el autor describa con detalle las condiciones de la línea base y tratamiento que ha aplicado; es decir, cuándo y cómo se produce la estabilización de la línea base y cuándo en la aplicación del tratamiento se empiezan a producir los efectos y su estabilización antes de interrumpir el tratamiento.

c) Duración de la fase de medida de la línea base y de los efectos del tratamiento. Por lo general, ambas fases deben durar el mismo tiempo, de otro modo se podría producir un desequilibrio que afectaría a las medidas, a los análisis estadísticos y a la interpretación. Sin embargo, motivos éticos pueden aconsejar hacer algo distinto, pero es el investigador el que debe decidir.

9.2. Tipo de diseños. Diseños A-B-A

Estos diseños se llevan a cabo en experimentos en los que intervienen muy pocos sujetos. La A expresa la medida o línea base y B representa la fase de tratamiento. Estos diseños recuerdan a los diseños de series temporales.

9.2.1. Diseño A-B

Es el diseño más simple. En una primera fase A, el investigador mide el comportamiento o la variable dependiente de los sujetos para establecer la línea base; a continuación, en la fase B, aplica el tratamiento y mide de nuevo el comportamiento. La comparación se hace entre ambas fases; es decir, las medidas obtenidas en la fase de ausencia de tratamiento y en la fase de aplicación del tratamiento. Este diseño, como podemos apreciar, tiene una baja validez interna. Si las medidas en la fase A son distintas de las de la fase B, podríamos concluir que los cambios que ocurran en la fase B (tratamiento) pueden deberse a otros factores distintos que han podido ocurrir en la fase de tratamiento o incluso en la fase de recogida de datos de la fase A (línea base).

9.2.2. Diseño A-B-A

En este diseño hay tres fases. En la primera fase se efectúa la medida de la línea base libre de tratamiento. En la segunda fase, la B, se aplica el tratamiento y se mide la variable dependiente. La tercera fase consiste en, una vez retirado el tratamiento, volver a la línea base y efectuar las correspondientes mediciones. Este diseño es parecido al anterior, pero al que hay que añadir la segunda condición de la línea base.

Este diseño tiene una buena validez interna. Si el comportamiento que estamos observando cambió en la fase del experimento, se podría concluir que los cambios observados han venido producidos por el tratamiento. El mayor inconveniente es la retirada del tratamiento, ya que si éste resulta positivo para los participantes, su retirada podría tener repercusiones negativas de tipo ético o académico para los implicados.

9.2.3. Diseño A-B-A-B

Este diseño tiene cuatro fases. En la primera fase hay una observación de la línea base A. En el segundo se introduce el tratamiento B. En la tercera se retira el tratamiento y se vuelve a la línea base (segunda línea base) y en la cuarta fase se reintroduce el tratamiento. Si las medidas del comportamiento varían en la segunda aplicación del tratamiento, el investigador puede tener el convencimiento de los efectos del tratamiento.

Una de las críticas más importantes que se ha hecho a este tipo de diseños ha sido su baja validez externa. Los resultados obtenidos no pueden generalizarse más allá del grupo objeto de investigación. Esta falta de validez externa puede ser compensada con estudios de réplica que incluyan variaciones en cuanto a sujetos, contextos, investigadores, etc. En cualquiera de los estudios originales o de réplica es importante la exhaustiva descripción de la fase de la línea base y de la fase de tratamiento, descripción de los sujetos, los instrumentos de medida, etc. Esta descripción va a ser el determinante del grado de generalización del estudio realizado.

En el análisis estadístico de este tipo de diseños muchos investigadores se limitan a informar de los datos directos y realizan análisis descriptivos. Los gráficos resultan de una gran utilidad. Son muy pocos los partidarios del uso de la estadística inferencial. Esto no significa que no se puedan utilizar técnicas como la *t* de Student para ver las diferencias que se producen entre fases. Otro tipo de análisis, y que sirven muy bien a este tipo de diseños, es el de series temporales. Muchos paquetes estadísticos incorporan un módulo de esta técnica, que aunque ofrece cierta complejidad, ha sido adaptada para su cálculo fácil e interpretación. Véase, por ejemplo, el paquete estadístico MYSTAT (Stegall y Hale, 1994).

10. LOS DISEÑOS CORRELACIONALES

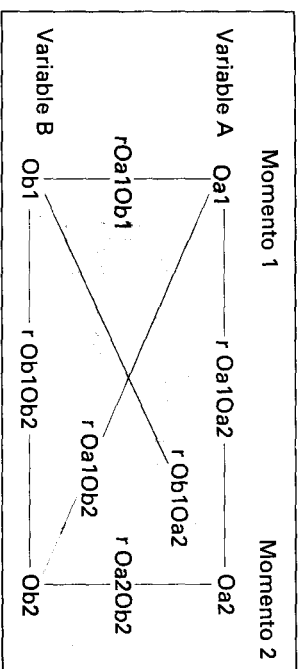
Los diseños anteriores tienen como característica básica que el investigador manipula la variable independiente o selecciona los valores de la variable independiente para ver sus efectos en la variable dependiente y, por tanto, establecer sus relaciones causales. Cuando el investigador no puede recurrir ni al control ni a la manipulación, le queda como recurso las técnicas correlacionales para inferir probables relaciones de causalidad entre las variables de estudio. Por ejemplo, cuando los datos disponibles provienen de estudios descriptivos o de observaciones, el investigador puede determinar la posible relación causal entre dos variables recurriendo a la técnica de correlación. Como muy bien sabemos la correlación en sí no es suficiente, aunque sí

necesaria para el establecimiento de relaciones causales entre variables. Sin embargo, con el desarrollo del panel de correlaciones cruzadas, técnica que utiliza varios coeficientes de correlación, es posible establecer algún grado de causalidad o direccionalidad entre dos variables (Huck, Cormier y Bounds, 1974).

Esta técnica de panel de correlaciones se basa en el factor tiempo en el sentido de que cuando un acontecimiento precede a otro (pero el reverso no puede ocurrir) el investigador puede establecer dos posibilidades:

- Que el acontecimiento A (aunque pueda haber otros) influye en el acontecimiento B.
- Que A y B están afectados por otro acontecimiento C.

La técnica de la correlación se usa precisamente para establecer la intensidad de la relación entre dos variables en momentos distintos y así poder inferir cierto grado de causalidad entre dichas variables. El diagrama que representa un panel de correlaciones es el que aparece a continuación:



Dos variables se han medido a un mismo grupo de sujetos en dos momentos diferentes, Momento 1 y Momento 2. Oa1 se refiere a la primera medida de una variable; Ob1 se refiere a la primera medida de la segunda variable. Ambas han sido medidas al mismo tiempo. La segunda serie de medidas de las mismas variables se hacen en un segundo momento y vienen representadas por Oa2 y Ob2. Estas cuatro medidas es lo que se denomina panel de datos. El primer panel viene determinado por Oa1-Ob1 y el segundo por Oa2-Ob2. A estos dos paneles se podría añadir otros paneles: Oa3-Ob3, etc.

Una vez efectuadas las observaciones o medidas, se hacen los cálculos de los coeficientes de correlación en un total de seis. Cada uno de estos coeficientes establece un tipo de relación entre las variables:

- Oa1-Ob1 y Oa2-Ob2 forman el primero y segundo paneles respectivamente.
- Oa1-Oa2 y Ob1-Ob2 expresan la consistencia de las variables.
- Oa1-Ob2 y Ob1-Oa2 son las correlaciones cruzadas y diferidas.

Las correlaciones que el investigador obtiene de estos cruces son las que le van a permitir, en relación con los otros coeficientes, determinar la causalidad entre las variables. Es decir, el cambio en una variable seguida por cambios en otra variable puede satisfacer la noción de causalidad.

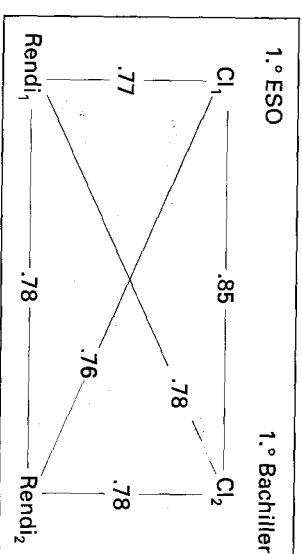
A partir de los coeficientes de correlación cruzados el investigador puede establecer cuatro hipótesis:

- H1:** Un incremento en A causará un incremento en B y un descenso en A causará un descenso en B.
- H2:** Un incremento en A causará un descenso en B y un descenso en A causará un incremento en B.
- H3:** Un incremento en B causará un incremento en A y un descenso en B causará un descenso en A.
- H4:** Un incremento en B causará un descenso en A y un descenso en B causará un incremento en A.

Este tipo de diseño facilita determinado tipo de investigaciones en el campo educativo, especialmente en aquellos casos donde no es posible la intervención, la manipulación de la variable independiente o la selección de niveles de la variable independiente. La validez interna de este diseño está limitado por algunos de los factores que al no ser controlados pueden afectar a dicha validez. Nos referimos a factores como la historia, la maduración, la ampliación de pruebas o la mortalidad.

Para el análisis estadístico de los datos, aparte de la técnica correlacional entre las cuatro medidas del panel, el investigador podría utilizar la *t* de Student para determinar si dos correlaciones son significativamente diferentes.

Un ejemplo de aplicación de este diseño podría hacerse al estudio de si la inteligencia es la causa del rendimiento académico. Para ello a un grupo de sujetos se le aplica una prueba de inteligencia y se obtienen sus calificaciones globales en dos momentos distintos de su vida escolar: 1.º de la ESO y 1.º de Bachillerato. Los resultados son los que aparecen en el panel que sigue (Huck, Cormier y Bounds, 1974).



11. EVALUACIÓN DE UN DISEÑO DE INVESTIGACIÓN. EJERCICIO PRÁCTICO

A continuación localiza en tu biblioteca los siguientes artículos y trabajos de investigación:

De la Orden, A. (1993). *Influencia de la evaluación del aprendizaje en la eficacia de la enseñanza*. Revista de Investigación Educativa, n.º 22 (2), pp 7-41.

Aguado Odina, M. T. (1993). *Análisis y valoración de programas pertenecientes al plan experimental para la reforma de la educación infantil*. Revista de Investigación Educativa, n.º 21 (1), pp 29-46.

Martínez Selva, J. M., y García Sánchez, F. A. (1992). Estudio de la atención fásica en niños con síndrome de Down. En Candel, I., y Turpín, A. (Eds.). *Síndrome de Down: Integración escolar y laboral*. Murcia: ASSIDO.

Responde a las siguientes preguntas:

1. ¿Qué participantes están implicados en el estudio? Describe sus características.
2. ¿En qué contexto o marco tuvo lugar la investigación? Describe el escenario: aula, laboratorio, etc.
3. ¿Cuál fue el tratamiento aplicado a los sujetos?
4. ¿Cuál fue la variable dependiente? ¿Qué instrumentos ha utilizado el autor para recoger los datos?
5. ¿Cuáles han sido los argumentos básicos utilizados por el autor o autores para establecer las relaciones causales?
6. ¿Cuándo tuvo lugar la investigación? ¿Cuánto tiempo duró el trabajo de campo o de laboratorio? ¿Cómo se han desarrollado todas las fases de la investigación? ¿Quedan descritas por el autor o autores?

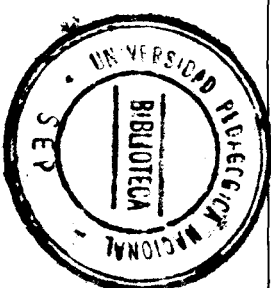
12. BIBLIOGRAFÍA

- Ato García, M. (1991). *Investigación en ciencias del comportamiento, I. Fundamentos*. Barcelona PPU-DM.
- Bisquerra, R. (1989). *Métodos de investigación educativa. Guía práctica*. Barcelona: CEAC.
- Borg, W. R., y Gall, M. D. (1979). *Educational Research*. New York: Longman.
- Bracht, G. H., y Glass, G. V. (1968). *The external validity of experiments*. American Educational Research Journal, 5, 437-74.
- Collyer, C. E., y Ems, J. T. (1987). *Analysis of variance: the Basic Designs*. Chicago: Nelson-Hall Pub.
- Del Rincón, D., y cols. (1992). *Investigación educativa. Fundamentos y metodología*. Barcelona: Labor.
- Fraenkel, J. R., y Wallen, N. E. (1993). *How to Design and Evaluate Research in Education*. New York: McGraw-Hill.
- Hernández Pina, F. (1995). *Bases metodológicas de la investigación educativa, I. Fundamentos*. Murcia: DM.
- Huck, S.; Cormier, W., y Bounds, W. (1974). *Reading statistics and Research*. New York: Harper y Row, Pub.
- Krathwohl, D. R. (1993). *Methods of Educational and Social Science Research. An integrated Approach*. New York: Longman.
- McMillan, J. H., y Schumacher, S. (1984). *Research in Education. A conceptual introduction*. Boston: Little, Brown y Co.
- Melá Navarro, J. L. (1990). *Introducción a la media y análisis de datos*. Edición Policopiadada. Valencia.
- Rosenthal, R. (1966). *Experimental effects in behavioral research*. New York: Appleton-Century. Crofts.

- Sprinthall, R. C.; Schumate, G. T., y Sirois, L. (1991). *Understanding Educational Research*. New Jersey: Prentice Hall.
- Stegall, J. W., y Hale, R. L. (1994). *MYSTAT for Windows*. Cambridge, MA: Course Technology, Inc.
- Tuckman, B. W. (1978). *Conducting Educational Research*. New York: Harcourt Brace Jovanovich, Inc.
- Verna, G. K., y Beard, R. M. (1981). *What is Educational Research*. Gran Bretaña: Gower.
- Welkowitz, J.; Ewen, R., y Cohen, J. (1976). *Estadística aplicada a las Ciencias de la Educación*. Madrid: Santillana.
- Wiersma, W. (1991). *Research Methods in Education*. London: Allyn y Bacon.
- Wynne, J. D. (1982). *Learning Statistics. A common-sense Approach*. New York: Macmillan Pub. Co., Inc.

LA INVESTIGACIÓN POR ENCUESTA

LEONOR BUENDÍA EISMAN
Universidad de Granada



UNIDAD SEAD 16A
MORELIA, MICH.

OBJETIVOS

1. Conceptualizar la metodología por encuesta y diferenciarla de otros modelos de investigación.
2. Desarrollar el proceso científico de la investigación por encuesta.
3. Diferenciar los muestreos representativos de los que no cumplen este requisito y elegir el más idóneo en función del propósito de la investigación.
4. Analizar los datos obtenidos con los cuestionarios y/o las entrevistas y elaborar conclusiones coherentes con los objetivos planteados.
5. Elaborar un informe de investigación

CONTENIDOS

1. Concepto y consideraciones metodológicas.
2. El proceso de la investigación por encuesta.
3. Diseños en la investigación por encuesta.
4. El muestreo.
5. Optimización del proceso.
6. Codificación y análisis de los datos
7. Elaboración de conclusiones
8. El informe de la investigación
9. Bibliografía

BIBLIOGRAFÍA BÁSICA

- Colás, P., y Buendía, L. (1994). *Investigación Educativa*. Sevilla: Alfar.
- Gómez, J. (1990). Metodología de encuesta por muestreo. En Arnau, J.; Anguera, M. T., y Gómez, J. *Metodología de la Investigación en Ciencias del Comportamiento*. Murcia: Secretariado de Publicaciones de la Universidad de Murcia.
- Hernández, F. (1995). *Introducción al proceso de investigación*. Barcelona: PPU.
- León, O., y Montero, I. (1993). *Diseño de investigaciones*. Madrid: McGraw-Hill.
- Rodríguez, J. (1991). *Métodos de muestreo*. Madrid: Centro de Investigaciones Sociológicas.

1. CONCEPTO Y CONSIDERACIONES METODOLÓGICAS

Con frecuencia, algunos autores identifican el método científico con el procedimiento seguido en la investigación experimental con las correspondientes condiciones de aleatorización, manipulación y control; dejando para la investigación descriptiva procedimientos subjetivos y métodos sistemáticos para describir la realidad. Para otros, el método científico, como método, no es nada más que sacar el mejor partido de la propia mente, sin barreras (Manheim, 1982: 79).

Participando de algunos elementos de ambos extremos, consideramos que cualquier procedimiento seguido para dar respuesta a un problema, caracterizado por el rigor y sistematicidad en el proceso, debe ser considerado científico.

En este capítulo vamos a tratar la encuesta como un método de investigación, capaz de dar respuesta a problemas tanto en términos descriptivos como de relación de variables, tras la recogida de información sistemática, según un diseño previamente establecido que asegure el rigor de la información obtenida.

Con esta información el investigador pretenderá: *a)* describir las condiciones de la naturaleza existentes, *b)* identificar normas o patrones con los que se puedan comparar dichas condiciones y *c)* determinar las relaciones que existen entre acontecimientos específicos.

Si situáramos en un continuo las diferentes metodologías de investigación, desde la metodología experimental, realizada en contextos artificiales con gran validez interna, hasta la metodología observacional, donde los fenómenos son estudiados en sus contextos naturales, con un control interno mínimo y buscando la descripción más completa posible del aspecto o sujeto investigado, la metodología por encuesta, considerada por algunos como metodología selectiva (Anguera, 1990), se situaría en un espacio intermedio. El control es fundamentalmente estadístico, y aunque las condiciones de realización son naturales, pierde el carácter ideográfico de la metodología observacional al tratar de representar un universo determinado, analizando la variabilidad existente mediante la selección adecuada de los sujetos que se caractericen por poseer, en distinto grado, la variable de interés.

Ninguna metodología prevalece sobre la otra por el mayor grado de científicidad. Su utilización va a depender del tipo de problema que se plantea el investigador y la necesidad de generalización de los resultados obtenidos.

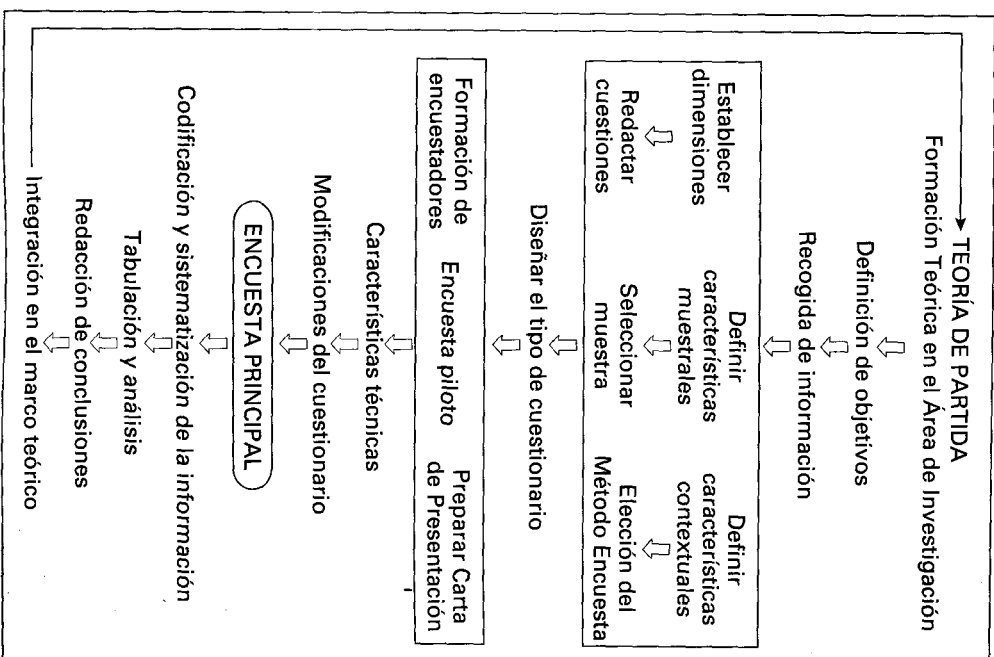
La utilización de la metodología por encuesta tiene ventajas en las siguientes situaciones:

- Cuando se quieren generalizar los resultados a una población definida, por ser mayor el número de elementos que forman la muestra que en otras metodologías, el desfase producido entre la muestra seleccionada y la población definida es menor.
- Es una importante alternativa cuando no es posible acceder a la observación directa por circunstancias contextuales o problemas económicos.
- Es la metodología más indicada para recoger opiniones, creencias o actitudes porque, si bien los encuestados pueden no decir lo que piensan realmente, al menos manifiestan lo que desean que el investigador sepa de ellos, por lo que es muy utilizada para obtener informaciones subjetivas de un gran número de sujetos.

Las ventajas de otras metodologías se presentan como limitaciones en ésta. La primera es la debilidad de los diseños para demostrar relaciones causales. En segundo lugar no tiene en cuenta los factores contextuales que pueden alterar la conducta del sujeto (Gómez, 1990). No obstante, es una metodología fundamental en los estudios sociales cuando el propósito es descriptivo y se necesitan muestras grandes para el estudio de algún aspecto de la población.

2. EL PROCESO DE LA INVESTIGACIÓN POR ENCUESTA

En el proceso de la investigación por encuesta se establecen tres fases de desarrollo: teórico-conceptual, metodológica y estadística-conceptual. El primer nivel incluye el planteamiento de objetivos y/o problemas e hipótesis de investigación. El segundo



nivel comienza en el momento de seleccionar la muestra y definir 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. *enriquecido con lo hallado.*

2.1. Formación teórica en el área de investigación

Como en cualquier otro proceso de investigación, antes de planificar el desarrollo o establecer los objetivos es necesario una formación teórica en los aspectos a investigar porque las dimensiones del cuestionario, e incluso cada una de las cuestiones en él planteadas, responden a un marco teórico en el que el investigador pretende profundizar o ver cómo se comportan determinadas variables en una población definida.

Difícilmente se puede planificar algo que no se conoce. Las obviedades sólo permiten acumular datos que no aportan nada o que en el mejor de los casos sólo sirven para justificar inversiones innecesarias. De una buena formación depende en gran medida el éxito de los resultados.

Para conseguir esta formación se puede recurrir tanto a las fuentes primarias como secundarias, y a investigaciones realizadas sobre el tema, bien por el método de encuesta o por otros métodos, que nos puedan ayudar a una mejor concreción de los objetivos y al diseño del cuestionario (Buendía, 1994a).

2.2. Planificación de la encuesta

La formulación de objetivos se deriva directamente del área problemática que pretendemos investigar y de los fines que perseguimos en esa parcela del conocimiento. Si la finalidad es conocer las características de una población, a partir de un estudio detallado de las variables de interés en un grupo o en diferentes subgrupos de sujetos, tendremos que planificar una encuesta descriptiva, que implica:

- Inclusión de datos personales y/o laborales que permitan realizar comparaciones entre las diferentes categorías.
- Trabajar con muestras representativas que permitan las extrapolaciones a la población definida.
- Recoger la información de muestras suficientemente grandes para minimizar el error, tanto en el tipo de distribución del fenómeno de estudio como en la intensidad de aparición, cuando el objetivo de la encuesta es extrapolarlo a un universo de alta variabilidad.

Si la finalidad es contrastar hipótesis o establecer relaciones causales, se planificarán encuestas explicativas que den razón de las relaciones entre las variables. Con estas encuestas se pretende responder «por qué» ocurren unas determinadas conductas o qué causa un determinado efecto.

La planificación de las encuestas explicativas implica:

- Haber acotado el campo de trabajo, tras un primer estudio exploratorio (encuesta de carácter descriptivo), y homogeneizar en lo posible las muestras para contrastar las hipótesis surgidas en la descripción.
- Al no poder manipular la variable independiente, la selección de la muestra tendrá que realizarse entre aquellos sujetos que ya la poseen naturalmente, estableciendo tantos subgrupos como niveles de dicha variable puedan aparecer.
- Utilizar técnicas de análisis adecuadas que permitan establecer las relaciones causales según los niveles de significación establecidos.

La planificación, tanto de la encuesta descriptiva como explicativa, responde a tres principios básicos: propósito que se persigue, población a la que va dirigida, y soporte material y humano con el que se cuenta.

El propósito de la encuesta no debe ser ni excesivamente ambicioso ni tan reducido que apenas permita la toma de decisiones.

Los objetivos generales o propósitos de la encuesta deben ser especificados claramente, porque de ellos va a depender toda la planificación.

La población debe estar bien definida atendiendo a criterios geográficos (lugar donde se realizará el estudio), demográficos (nivel socioeconómico, edad, sexo, etcétera) y temporales (momento en el que se va a recoger la información). Si la población definida es pequeña, entrarán a ser objeto de estudio todos los individuos pertenecientes a ella. Si, por el contrario, la población es muy numerosa, se precisará realizar un muestreo representativo que permita generalizar los resultados obtenidos con la muestra a la población de la cual fue extraída.

El soporte material y humano condicionan en gran medida el desarrollo de la encuesta. Hay veces que por recorte presupuestario hay que recurrir al muestreo, cuando el interés del tema y las características del estudio exigirían realizarlo con toda la población. El coste económico de entrevistadores, desplazamientos, codificación y corrección de cuestionarios, etc., deben ser estudiados en la planificación para ajustar la ampliación del estudio al presupuesto disponible.

2.3. Elaboración de instrumentos de recogida de datos

La investigación por encuesta debe traducir las variables empíricas sobre las que se desea obtener información en preguntas concretas sobre la realidad social a investigar, capaces de suscitar respuestas únicas y claras (Sierra Bravo, 1985: 272).

Esto se consigue a través del cuestionario y de la entrevista. Simon (1978) apuntaba otra forma de obtener datos, que es el análisis de materiales existentes.

2.3.1. El cuestionario

El cuestionario ha sido la técnica de recogida de datos más utilizada en la investigación por encuesta. Con él se pretende conocer lo que hacen, opinan o piensan los

encuestados mediante preguntas realizadas por escrito y que pueden ser respondidas sin la presencia del encuestador.

Esto, en encuestas sociales, es de una gran complejidad por la dificultad que entraña expresar en indicadores concretos los aspectos de la realidad que interesan investigar. Esta dificultad puede disminuirse si se cuenta con un equipo de personas para su elaboración y, antes de comenzar realmente a construirlo, se dedican algunas jornadas a delimitar el campo de estudio y a definir claramente los indicadores que interesan en dicho campo.

La labor se facilita si se opta por la realización de una preencuesta abierta para recoger opiniones que serán la guía para la formulación posterior.

En esta preencuesta pueden utilizarse entrevistas individuales e incluso documentación sobre el tema o aportaciones de otras encuestas similares. Toda esta información se someterá a juicio crítico sobre su procedencia o no de inclusión en la construcción del instrumento.

A) Elaboración de un cuestionario

Cuando se elabora el cuestionario se deberán realizar suficientes preguntas para que queden reflejados todos los aspectos importantes según el propósito de la encuesta. En el planteamiento de las diferentes fases seguiremos el trabajo de Buendía (1994b: 207-211).

Se especificarán:

1. Las cuestiones con datos de identificación y clasificación. Por ejemplo: edad, nivel de estudios, estado civil, etc.
2. El tipo de preguntas más adecuado.

Cuando el cuestionario es autoadministrado, son preferibles las preguntas cerradas, por la facilidad que proporcionan este tipo de preguntas para la interpretación de las respuestas.

Si se realizan preguntas abiertas, deben ser de respuesta breve, porque las que exigen respuestas largas no suelen ser contestadas, bien por no perder tiempo o por temor a no saber expresarse, perdiéndose un importante número de respuestas. A veces es preferible dejar un espacio al final que permita hacer comentarios al que lo dese.

3. El orden de las preguntas y la disposición en el cuestionario.

En cuanto a la secuencia de las cuestiones es conveniente presentar primero las más interesantes, procurando que no sean difíciles o embarazosas. Se debe evitar colocar juntas preguntas que la respuesta a una de ellas pueda influir en el sentido de la respuesta a la otra.

4. Si el cuestionario trata varios temas, las preguntas sobre un mismo tema deben ir juntas, pues de lo contrario podrían desorientar al sujeto.

El número de preguntas, de forma que contenga las necesarias pero ninguna más.

5. Determinar las preguntas de «alivio», más o menos intrascendentes, que deben introducirse para evitar la monotonía y el cansancio, o para facilitar el cambio de tema.

6. Redacción de las preguntas. Deben ser lo más claras y sencillas posible, procurando que el lenguaje sea el apropiado para las personas a las que va dirigido.
7. Determinar los aspectos formales tales como:
 - Clase, color y tipo de impresión.
 - Espacios destinados a la respuesta.
 - Tipo de codificación.
 - Márgenes, etc.
8. Redactar los escritos que deben acompañar al cuestionario:
 - Carta dirigida al encuestado, en la que se solicita su colaboración, ofreciéndole información del interés de la investigación y agradeciéndole su participación.
 - Instrucciones para cumplir el cuestionario.

Las normas, que deben acompañar a un cuestionario para su cumplimentación, tienen tanta importancia como el cuestionario, pues permiten que se puedan normalizar las respuestas recogidas y evitar los sesgos derivados de la interpretación subjetiva.

CUESTIONARIO

- Datos de identificación y clasificación

- Preguntas

- * Tipo de preguntas
- * Orden y disposición
- * Número de preguntas

- Aspectos formales

- Normas de cumplimiento

B) Realización del cuestionario

Los cuestionarios se pueden entregar directamente, o mediante correo. Cuando son entregados directamente, el encuestador puede explicar los objetivos del estudio, y permite al encuestado una primera lectura para que pregunte las dudas que puedan surgirle sobre la investigación, o aclaraciones a las preguntas del cuestionario.

Es conveniente fijar una fecha para pasar a recogerlo directamente. Esto implica más a los encuestados y el número de encuestas extraviadas, o no contestadas, suele disminuir significativamente.

El cuestionario entregado personalmente es de gran utilidad cuando lo tiene que contestar un grupo de personas que conviven juntos o trabajan en el mismo lugar, y con poco gasto de tiempo y dinero se recoge una información precisa. Por ejemplo, si el trabajo se va a realizar en colegios públicos, es conveniente entregarlos a la hora del recreo y se gana en tiempo y validez de la información.

Otra forma de entrega y recogida de cuestionarios es a través del correo. Sin duda es el que más problemas ha presentado a los investigadores por su baja tasa de respuesta. De hecho, Haberlein y Baumgartner (1978) apuntaban que las respuestas a un cuestionario variaban de un 46 por 100 cuando el encuestador realiza sólo un envío, a un 84 por 100 cuando repitían los envíos tres veces más, recordando la necesidad de recoger las respuestas. Estos autores insisten en la necesidad de realizar un seguimiento de los cuestionarios como la mejor manera de aumentar las tasas de respuesta. Otros factores influyentes son: la longitud del cuestionario, la claridad de comprensión de las preguntas, el tema tratado, la credibilidad de la persona que lo dirige o la institución que lo respalda, la personalización de la carta de presentación, y la posibilidad de incentivar económicamente a las personas que lo realizan en concepto de ayuda a la investigación, así como la disponibilidad anónima de la información que se recoge, una vez finalizada la investigación, por las personas que lo soliciten.

Dillman (1983) recomienda seguir los siguientes pasos para conseguir un mayor éxito en las respuestas:

1. La carta de presentación deberá incluir la utilidad social de la investigación, importancia de la respuesta personal y persona que debe rellenar el cuestionario.
2. Esta carta deberá llevar el membrete de la institución que respalda el estudio, con fecha y firma del responsable de la investigación.
3. Si el cuestionario se marca con un número de identificación, éste debe incluirse en la carta de presentación.
4. Esta carta, el cuestionario, y un sobre con el sello y la dirección del responsable de la encuesta se incluyen en otro sobre con la dirección personal del destinatario.
5. En breve espacio de tiempo, aproximadamente una semana, se envía una carta de recordatorio. Tres semanas después se puede volver a recordar a los que no han contestado, enviando un nuevo cuestionario, si las tasas de respuesta no han llegado al nivel deseado.

C) Ventajas y limitaciones

Una de las mayores ventajas de este instrumento es que no se necesitan personas preparadas para la recogida de información; la mayoría de las veces ni siquiera se realizan entregas personales, enviándose por correo el cuestionario con una simple carta de presentación.

Las respuestas pueden ser más meditadas, al no exigirse tiempo en la realización, además de existir mayor libertad de expresión al no encontrarse coaccionado el encuestado por la presencia del entrevistador.

Su principal inconveniente y uno de los mayores problemas en la utilización de esta técnica de encuesta es el elevado número que se pierde cuando son enviados por correo; aproximadamente un 30 por 100. Esto plantea siempre al investigador la duda de si los que no contestan son diferentes de los que contestan, y si es así, la investigación es sesgada y ha perdido su representatividad. La única solución es la repetición de cartas o visitas solicitando la colaboración; aunque al ser anónimo habrá que especificar en la carta la necesidad de volver a insistir, a la vez que se piden disculpas a los que hubieran respondido.

Otro aspecto a considerar en el cuestionario es que las respuestas pueden estar afectadas por el cambio de humor o estado general. A esto hay que unir que sólo es adecuado para determinadas personas. Contestarlo implica no sólo cierto nivel de comprensión y expresión, sino de dominio de las «técnicas de respuesta» (Pérez Juste, 1984: 108), por lo que muchos cuestionarios se devuelven incompletos, interpretándose como opinión del encuestado lo que es realmente una falta de adecuación del instrumento a la población que va dirigido.

En líneas generales, el cuestionario por correo no está indicado cuando la población a la que va dirigido es anciana o personas de bajo nivel cultural. En estos casos es preferible realizar entrevistas.

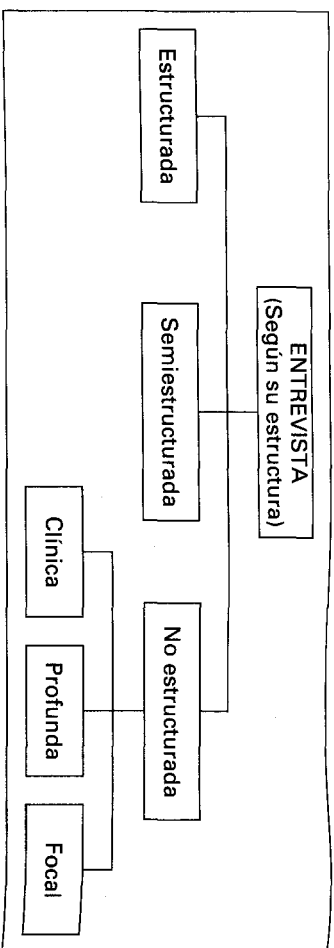
2.3.2. La entrevista

La segunda técnica de recogida de datos es la entrevista.

Ésta consiste en la recogida de información a través de un proceso de comunicación, en el transcurso del cual el entrevistado responde a cuestiones, previamente diseñadas en función de las dimensiones que se pretenden estudiar, planteadas por el entrevistador.

Consideramos que si bien la entrevista no estructurada es la técnica por excelencia de la investigación cualitativa (Denzin y Lincoln, 1994), en los estudios de encuesta es la entrevista estructurada o muy estructurada la que ocupa este papel. Ambas entrevistas, estructurada y no estructurada, son técnicas al servicio de lógicas de investigación diferentes (justificación versus descubrimiento) y por tanto con procesos y fines igualmente distintos.

A continuación se exponen las diferencias entre una y otra (Ruiz 1996: 169), como técnicas de recogida de información.



A) Entrevista estructurada

Se realiza partiendo de un cuestionario previamente elaborado y cuya principal característica es la inflexibilidad, tanto en las cuestiones a plantear al entrevistado como en el orden y presentación de las preguntas. Tiene la ventaja de reducir los sesgos del entrevistador, pero impide indagar al entrevistador con preguntas complementarias sobre cuestiones que pueden ser de interés, ni pedir aclaraciones sobre respuestas ambiguas. A veces se recurre a la entrevista semiestructurada que es más flexible y abierta. El entrevistador tiene la libertad para alterar el orden y la forma de preguntar así como el número de preguntas a realizar. Se dispone de un guión base que puede modificarse por intereses de la entrevista, aunque manteniéndose el objetivo para el cual fue preparado y los diversos puntos sobre los que debe obtenerse información.

Requiere entrevistadores muy entrenados y que a la vez conozcan ampliamente el atributo que se estudia. Sus características son:

- Pretende explicar más que comprender.
- Busca minimizar los errores.
- Adopta el formato estímulo/respuesta, suponiendo que a una respuesta correcta el entrevistado contesta con la verdad.
- Obtiene con frecuencia respuestas racionales, pero pasa por alto la dimensión emocional.

A.1) El entrevistador

- Formula una serie de preguntas con una serie de respuestas prefijadas entre las que elegir.
- Controla el ritmo de la entrevista siguiendo un patrón estandarizado y directo.
- No da explicaciones largas del trabajo.
- No altera el orden ni la formulación de las preguntas.
- No permite interrupciones ni intervención de otras personas en las respuestas.
- No expresa su opinión personal en favor o en contra. Su papel es el de un rol neutral.
- No interpreta el sentido de las preguntas, sólo da las explicaciones previas.

- Nunca improvisa el contenido o forma de las preguntas.
- Establece una «relación equilibrada» que implica familiaridad al mismo tiempo que impersonalidad.
- Adopta el estilo de oyente interesado, pero no evalúa las respuestas dadas.

A.2) El entrevistado

- Todos los entrevistados reciben el mismo paquete de preguntas. Las escuchan en el mismo orden y formato.

A.3) Las respuestas

- Son cerradas al cuadro de categorías preestablecido.
- Grabadas conforme al sistema de codificación previamente acordado.

B) La entrevista no estructurada

Es la más utilizada en los estudios exploratorios. La flexibilidad es total. Sólo se determina a priori el tema de la entrevista. El entrevistador tiene libertad para formular el contenido, el número y la secuencia de las preguntas. Esta entrevista exige una alta preparación del entrevistador, pues, en la medida que avanza, exige ir creando las cuestiones que interesan conocer, sin olvidar la situación global para la que está siendo realizada. Sus principales características son:

- Pretende comprender más que explicar.
- Busca maximizar el significado.
- Adopta el formato de estímulo/respuesta sin esperar la respuesta objetivamente verdadera, sino subjetivamente sincera.
- Obtiene con frecuencia respuestas emocionales, pasando por alto la racionalidad.

B.1) El entrevistador

- Formula preguntas sin esquema fijo de categorías de respuesta.
- Controla el ritmo de la entrevista en función de las respuestas del entrevistado.
- Explica el objetivo y motivación del estudio.
- Altera con frecuencia el orden y forma de las preguntas, añadiendo nuevas si es preciso.
- Permite interrupciones e intervención de terceros si es conveniente.
- Si es requerido, no oculta sus sentimientos ni juicios de valor.
- Explica cuanto haga falta el sentido de las preguntas.
- Con frecuencia improvisa el contenido y la forma de las preguntas.
- Establece una relación equilibrada entre familiaridad y profesionalidad.
- Adopta el estilo del oyente interesado pero no evalúa las respuestas.

B.2) El entrevistado

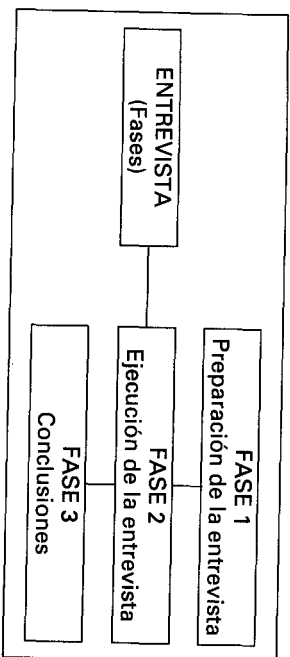
- Cada entrevistado recibe su propio conjunto de preguntas.
- El orden y formato puede diferir de uno a otro.

B.3) Las respuestas

- Son abiertas por definición, sin categorías de respuestas preestablecidas.
- Grabadas conforme a un sistema de codificación flexible y abierto a cambios en todo momento.

C) Fases en la elaboración de una entrevista

En el proceso de una entrevista hay que contemplar las siguientes fases:

**C.1) Preparación de la entrevista**

La preparación de la entrevista va a determinar en gran medida el éxito de ésta. El entrevistador debe procurar una buena impresión. Las tres cualidades básicas que debe ofrecer al entrevistado son: aceptación, comprensión y sinceridad.

El lugar donde se realiza debe estar alejado del ruido y de miradas indiscretas, para que haya una mayor y más fácil comunicación entre ambos. Se procurará un ambiente acogedor y relajante.

El entrevistado debe conocer, si él no ha pedido la entrevista, el objetivo de la misma y los motivos que justifican su realización. Debe comunicársele el carácter estrictamente confidencial y anónimo de la información que facilite.

Toda entrevista exige un guión donde se determinen los objetivos y aspectos a tratar, y si la entrevista es estructurada, cada aspecto incluirá una o varias preguntas que el entrevistador debe conocer perfectamente para no dar la impresión de improvisación durante el desarrollo de la misma.

C.2) Ejecución de la entrevista

En la primera toma de contacto, el entrevistador, si no fuera conocido por el entrevistado, debe presentarse y explicar claramente los extremos anteriormente comentados.

La encuesta puede tenerla el entrevistador en la mano, pero las preguntas deben realizarse en la forma más natural posible, evitando dar la impresión de un interrogatorio.

Entre pregunta y pregunta no debe mediar tiempo, sino que su desarrollo debe ser como el de una conversación cualquiera, introduciendo frases del tipo: «bien», «de acuerdo», «le escucho», y siempre prestando más atención al entrevistado que a la encuesta que hay en la mano.

Si la entrevista fuera hecha a varias personas, debe procurarse no introducir cambios en el momento de las preguntas, evitar el sesgo de las opiniones personales del entrevistador, y las variaciones de significado por el cambio de palabras o giros en las preguntas.

Una entrevista estructurada, con un cuestionario previo, debe mantener el orden de las preguntas; generalmente la disposición de éstas guarda una estrecha relación con el efecto de contagio de las respuestas.

El entrevistador no debe dar nunca su opinión, y por consiguiente no se admite ningún tipo de discusión ni de justificación.

C.3) Conclusiones

Si las respuestas no se anotan en el momento en que se realiza la entrevista, se presentan dos clases de problemas: El olvido puede jugar un papel negativo en los resultados de la encuesta y en segundo lugar puede producirse cierta distorsión en el recuerdo por los elementos subjetivos que se proyectan en la reproducción de las situaciones.

Es por esto recomendable ir tomando las notas precisas, y si lo permite el entrevistado debería grabarse, para ganar en veracidad y fiabilidad.

El final de una entrevista debe tener un clima de cordialidad para que el entrevistado no acabe con sensación de desconfianza e irritabilidad por el desarrollo de la misma.

D) La entrevista por teléfono

Otra forma de recoger información es a través de la entrevista por teléfono.

Es especialmente útil cuando se abordan temas delicados que exigen impersonalización, recogiendo aspectos o comportamientos negativos que de otra forma serían difíciles de obtener. Este aspecto ha sido debatido, con opiniones contradictorias tanto en las tasas de respuesta como en la veracidad de las informaciones. Para Fowler (1988) es adecuado por la impersonalidad que acompaña la entrevista por teléfono en los temas de delicada respuesta; sin embargo, Groves (1979) constató que los entrevistados prefieren el procedimiento personal al telefónico e incluso hay mayores tasas de no respuesta o de dar respuestas falsas.

La selección de los encuestados se puede hacer a través del listín telefónico o a través de la aleatorización de la secuencia de dígitos. Este segundo método consiste en la selección de llamadas telefónicas con la ayuda de un generador de números aleatorios. En cualquier caso la selección recae sobre un hogar, con un número de

personas no determinado en la selección. La entrevista se realiza sobre un miembro de ese hogar elegido de forma aleatoria.

Últimamente las técnicas de entrevista telefónica se han visto mejoradas gracias al uso del ordenador. La entrevista telefónica asistida por ordenador: CATI (Computer Assisted Telephone Interviewing), ha supuesto una gran reducción de tiempo y preparación de expertos en muestreo. El ordenador selecciona y controla la muestra, recoge las informaciones conforme se van produciendo y realiza automáticamente los pasos de codificación y edición (Shure y Meeker, 1978). Es por tanto una técnica de recogida de información de gran futuro puesto que el ahorro de tiempo y personal es enorme y se evitan muchos de los problemas planteados con la selección a través del listín telefónico o la necesidad de muestras pequeñas que exige la aleatorización de la secuencia de dígitos.

E) Ventajas y limitaciones

Este instrumento, dependiendo de los fines para los que se utiliza, tiene ventajas y limitaciones. Entre las ventajas se pueden destacar las siguientes:

- Se puede recoger información de personas con bajo nivel cultural e incluso analfabetas.
- Permite conocer, además de las respuestas, el estado de ánimo y el ambiente del entrevistado.
- Es posible obtener mayor información que con los cuestionarios cuando se trata de cuestiones comprometidas.
- Existe menor riesgo de pérdida que con los cuestionarios; generalmente se accede a las entrevistas, y aunque existe un porcentaje de fallidas, siempre es menor que la pérdida de los cuestionarios.

Entre las limitaciones encontramos:

- El sesgo que introduce el entrevistador, por las diversas influencias que puede ejercer durante el desarrollo de la entrevista.
- No mantiene de forma tan evidente el anonimato de las respuestas como el cuestionario.
- Es necesario que los entrevistadores sean expertos para que los datos recogidos sean lo más fiables posibles.

2.3.3. El análisis secundario

La utilización del material existente (Kiecolt y Nathan, 1985) para nuevos fines o apoyo de nuevos objetivos de investigación ha sido bastante frecuente en las últimas décadas (Hyman, 1972), gracias a los medios técnicos de almacenaje y a la facilidad del investigador para acceder a los archivos de datos. No se trata, pues, de recurrir a los meta-análisis sobre los temas, sino a los datos primarios, que siguen siendo la unidad de análisis, aunque con fines distintos o para comparaciones de hipótesis en el tiempo.

En el análisis secundario se distinguen tres tareas claves (Gómez, 1990):

- a) Recolección de datos.
- b) Identificación de sesgos y errores.
- c) Elección de indicadores e índices.

La primera fase, recolección de datos, se puede realizar sobre un solo fichero o sobre la combinación y posterior selección de datos de diferentes ficheros. Hyman (1972) describe diferentes tipos de datos en función de la selección realizada.

Éstos son:

- Análisis secundario de los datos de una sola encuesta por muestreo.
- Agrupación de diversas encuestas (con indicadores equivalentes) en un único conjunto de datos.
- Replicación interna de múltiples estudios, con indicadores equivalentes.
- Replicación externa de una encuesta previa por otra, que a su vez puede ser replicada por otra más, hasta formar una serie de tiempo de encuestas relacionadas.
- Replicación intra-encuesta, en la que diferentes indicadores de una misma variable son analizados por separado.
- Reducción de la muestra, de modo que los sujetos que no se adecuen a cierto criterio sean excluidos del análisis.
- Análisis combinado de múltiples encuestas para sintetizar características generales de determinados grupos, a partir de características más específicas.

En la segunda fase, se trata de identificar los errores cometidos en la selección de los datos, tanto en el momento de la selección como en el posterior registro y codificación.

En la tercera fase se seleccionan indicadores precisos para el tema de estudio, teniendo presente que se parte de estudios que fueron concebidos para propósitos posiblemente diferentes.

La principal ventaja de la obtención de datos secundarios es la cantidad de tiempo y dinero que se ahorra, pero hay que tener la precaución de utilizar archivos de datos cuyas unidades de análisis sean las adecuadas al nuevo estudio y que los datos no estén obsoletos para el momento en el que se necesitan.

3. DISEÑOS EN LA INVESTIGACIÓN POR ENCUESTA

Entendemos por diseño un plan estructurado de acción que, en función de unos objetivos básicos, está orientado a la obtención de datos relevantes a los problemas y cuestiones planteadas. Entendido así, el diseño puede hallarse en cualquier tipo de investigación; en todo caso se pide una coherencia lógica en las etapas previstas para alcanzar los objetivos de la investigación y una justificación científica de la validez de la estrategia elegida. Esta flexibilidad en la concepción del diseño no contraviene la rigurosidad que el término tiene en un contexto explícitamente experimental (Tejedor, 1984).

Los elementos básicos que pueden determinar la elección de un tipo u otro de diseño están en relación directa con los objetivos específicos del estudio y la naturaleza del mismo.

Los objetivos en este tipo de estudios pueden oscilar desde los puramente descriptivos hasta la explicación de relaciones causales, pasando por la identificación de regularidades en el conjunto de los datos, mediante el análisis de las interrelaciones de las variables, con el objeto de obtener tipologías o taxonomías de clasificación, o de identificar constructos o variables latentes.

En lo que se refiere a la clasificación de los diseños en la metodología por encuesta, dada la ausencia de manipulación directa de las variables, la estructura diferencial de los diseños viene determinada por aspectos tales como: las variables seleccionadas, el momento y número de medidas y la selección de los grupos. En función de la dimensión temporal en la recogida de los datos, diferenciamos entre diseños transversales y diseños longitudinales.

3.1. Diseños transversales

El diseño transversal es el más utilizado en la investigación por encuesta. Los datos se recogen, sobre uno o más grupos de sujetos, en un solo momento temporal; se trata del estudio en un determinado corte puntual en el tiempo, en el que se obtienen las medidas a tratar (Dwyer, 1983). El término *transversal* más que definir un diseño concreto define una estrategia que está implicada en una gran variedad de diseños. Sin embargo, a pesar de esta posible diversidad, su carácter transversal les lleva a compartir ciertos supuestos, objetivos y aspectos metodológicos, que hace pertinente su tratamiento unificado (al menos dentro de la investigación por encuesta).

Ejemplo de diseños transversales

En una investigación realizada para conocer la actitud hacia la reforma del sistema educativo de los profesores de EGB y Formación Profesional (Buenfía, 1991), en uno de los ítems se deseaba conocer la relación existente entre el tipo de centro y el grado de acuerdo con una determinada opinión sobre la reforma.

Actitud						
Tipo de centro	Totales	1	2	3	4	5
Públicos	65	12	20	15	18	5
Concertados	68	19	19	6	10	19
Privados	12	3	4	1	3	1

(1: totalmente de acuerdo; 2: de acuerdo; 3: indiferente; 4: en desacuerdo; 5: en total desacuerdo)

La selección específica de los grupos de sujetos a incorporar en el estudio es otro de los elementos bajo el control del investigador que contribuye a determinar la estructura básica de estos diseños. A través de ellos se pueden realizar estudios comparativos clasificables en dos categorías generales: los estudios de grupos naturales pertenecientes a un mismo entorno cultural (diseño intracultural) y los estudios de comparación entre diferentes culturas o naciones (diseño transcultural).

3.2. Diseños longitudinales

El término de diseños longitudinales, al igual que decíamos con los transversales, es amplio y aplicable a diferentes tipos de diseños. El denominador común de todos es la obtención de datos de un grupo determinado en diferentes puntos de tiempo, obteniéndose medidas repetidas a lo largo de un determinado parámetro temporal. Por ejemplo, el seguimiento de los votos hacia líderes políticos a lo largo de las campañas electorales.

Entre los estudios de encuesta que comparten la estrategia longitudinal establecemos una primera diferenciación en función de la amplitud de la secuencia observada, distinguiendo entre:

- Diseños de series temporales.*
- Diseños de panel.*
- Diseño de tendencias.*
- Diseño de cohorte.*

a) Diseños de series temporales

En los diseños de series temporales la amplitud de la secuencia observada es mucho más amplia que en los diseños de panel. Cuando la finalidad de la serie es fundamentalmente descriptiva o predictiva, no existe intervención específica que provoque «interrupción» en la serie de medidas registradas, siendo éste el diseño de series temporales más propio de la investigación por encuesta. El desarrollo de estos diseños puede verse en la página 102, en el apartado de diseños cuasi experimentales.

b) Diseños de panel

El diseño de «panel» implica la recogida de datos de un mismo grupo de sujetos, respecto a las mismas variables, al menos en dos momentos.

Este diseño tiene la dificultad de conseguir a los mismos sujetos en sucesivas mediciones. La mortalidad en las encuestas de panel suele ser alta debido al cambio de domicilio, o incluso a los cambios de opiniones respecto a su colaboración.

Sin embargo, el mayor problema en este diseño es la reactividad de los sujetos, debido a la sensibilización de los entrevistados hacia los temas de la encuesta, al haber sido ya encuestados, sobre los mismos temas, en un periodo anterior.

c) Diseño de tendencias

Estos diseños se utilizan cuando el número de puntos de medidas aumentan, de forma que permiten el análisis de tendencias. En este sentido constituye un punto intermedio entre el diseño de panel y el de series temporales.

Se caracteriza por realizar comparaciones de datos transversales en dos o más momentos temporales con el objeto de conocer los cambios en una variable, o en las relaciones entre variables, de una población. Cada vez que se realiza una medición se muestra nuevamente la población, por lo que los cambios observados no pueden ser atribuidos a cambios en los individuos, puesto que la población entrevistada no es la misma.

Este diseño evita la dificultad de localizar a los sujetos muestreados previamente, el problema de la mortalidad de la muestra y la reactividad de los estudios de panel (Gómez 1990: 272).

Por ejemplo si la encuesta planteada anteriormente para conocer la actitud de los profesores hacia la reforma del sistema educativo, realizada en el año 1988, se repite en el 1992 y posteriormente en el 1997, cuando está totalmente implantada la LOGSE.

d) Diseño de cohorte

Las dificultades para el estudio separado del efecto de edad y el efecto de «cohorte» (generación por fecha de nacimiento o acontecimiento histórico) en los datos de un diseño longitudinal simple impulsan una estrategia alternativa que permite la descripción simultánea de los cambios relacionados con una y otra variable (Schaie, 1965) a través de los diseños de cohorte.

El término «cohorte» se refiere generalmente a sujetos de la misma edad, aunque a veces pueda extenderse a grupos de sujetos que han vivido un importante período de tiempo con algún interés social o histórico. Por ejemplo, el estudio de los intereses ecológicos de los adolescentes de los años sesenta, o bien el estudio de los valores en el año 1977 (año de aprobación de la constitución) de los diferentes partidos políticos y su evolución posterior.

Ejemplo de diseño de cohorte

Grupo de adolescentes que tenían 14 años en el año 1960: Valores dominantes y evolución en el tiempo.

Cohorte	Ecológ.	Relig.	Social.	Polít.	Económ.	Otros
1960						
1970						
1980						
1990						

Están especialmente indicados cuando se desea conocer el cambio social, económico, o político, y en los estudios sobre desarrollo evolutivo.

No obstante, a veces se pueden realizar combinaciones de los diferentes diseños, porque así lo requiera la investigación, realizando extensiones de los diseños de panel, de tendencias o de cohorte.

4. EL MUESTREO

Al conocimiento de una realidad no siempre es posible acceder a través de todos los casos posibles que comparten características comunes; a veces sólo podemos seleccionar una parte de ese universo que denominamos muestra.

El concepto de muestra y población dependen de la definición que previamente se haya realizado en función del objetivo de la encuesta. Si el objetivo es conocer las formas de vida de los universitarios andaluces, la población la forman todos los alumnos matriculados en las diferentes universidades de Andalucía. Si el objetivo es conocer la forma de vida de los universitarios europeos y se realiza un muestreo de comunidades, es posible que lo que fue definido como población pase a ser parte de una muestra en una investigación diferente. La muestra es, por tanto, un subconjunto representativo de la población total, que incluye todas las características extrapoblacionales que desean ser conocidas con la información recogida y posteriormente tengan la misma probabilidad de ser muestreados; de no ser así, se obtendrán muestras sesgadas o tendenciosas. Por ejemplo, si deseamos conocer las opiniones de los jóvenes de 20 años hacia el deporte y para esto preguntamos sólo a los alumnos de Educación Física, los resultados no representan el sentir de los jóvenes españoles de esa edad, sino sólo los de ese grupo en concreto.

El concepto básico de la investigación por encuesta es el de muestra representativa. Esto quiere decir que el subconjunto de elementos que la forman comparten las características de la población a la cual se pretende generalizar los resultados.

Cuando se realiza una investigación por encuesta, obtener una muestra representativa de la población supone conseguir la información requerida con un importante ahorro de recursos económicos, de personal y de tiempo, y con posibilidad de extrapolar los resultados a toda la población.

Para conocer si una muestra es representativa han de tenerse en cuenta tres factores: primero, el proceso de selección; segundo, la estructura de la muestra, y, por último, el tamaño de la muestra.

El proceso de selección determina en gran medida la representatividad de la muestra. Las muestras probabilísticas se caracterizan porque en ellas cada elemento de la población tiene una probabilidad conocida y no nula de ser seleccionado (Kish, 1972). Existen cuatro métodos de muestreo probabilístico:

- Muestreo aleatorio simple.
- Muestreo aleatorio sistemático.
- Muestreo aleatorio estratificado.
- Muestreo aleatorio por conglomerados.

Los tres primeros se denominan monoetápicos por elegirse los elementos de la muestra en una sola etapa. El último es polietápico puesto que los elementos de la muestra se eligen en dos etapas sucesivas; en la primera se seleccionan los conglomerados de elementos de los que posteriormente se obtendrá la muestra.

La elección de un método u otro depende de las características de la investigación y de la viabilidad de la misma. Hay veces que sería conveniente elegir el método de muestreo aleatorio simple, por la precisión y características del trabajo, pero resulta inviable acceder a todos los miembros de la población sobre la que se desea realizar el muestreo, por lo que se hace necesario recurrir a otro, quizá menos preciso pero más viable.

Los métodos no son excluyentes entre sí. De hecho, en el muestreo estratificado se recurre a elegir de forma aleatoria los elementos que constituyen cada estrato. Así pues, no tienen por qué utilizarse de forma aislada, sino que con frecuencia se aplican simultáneamente en la selección de la muestra.

La estructura de la muestra es otro aspecto importante a tener en cuenta para evitar los sesgos en las estimaciones. Es necesario concretar perfectamente la población que va a ser objeto de estudio, para conocer las características que la componen y poder definir el marco del cual va a ser extraída la muestra. En la medida que más conocido es el marco, menos errores se cometerán en el diseño de selección. Por ejemplo, si para realizar un muestreo de la forma de vida de los vecinos de una comunidad se utiliza un listado de hace dos o más años, podemos encontrarnos con que la información recogida no refleja en absoluto la vida de ese barrio, puesto que si es una zona de fuerte acogida de emigrantes en los últimos años, estas personas no formarían parte de la población de estudio, con el consiguiente sesgo que se produce por desconocimiento del marco actual.

En cuanto al tamaño de la muestra, es una de las mayores preocupaciones de los investigadores, y aunque ciertamente no carece de importancia, sin embargo depende del propósito del estudio y de la homogeneidad de las características de la población, su mayor o menor amplitud. En líneas generales, se denomina una muestra suficiente cuando es mayor de treinta, pero, ciertamente, ésta puede no ser significativa en una población con características muy heterogéneas, o excesivamente amplia cuando la característica de estudio tiene poca presencia en la población. Por ejemplo, si deseamos conocer las creencias de los alumnos ciegos sobre sus posibilidades de inserción laboral en las diferentes comunidades autónomas, los elementos de la muestra en cada comunidad serán proporcionales al número de ciegos en las diferentes comunidades, y en algunos casos treinta puede ser excesivamente grande. En la página siguiente volveremos sobre el tema cuando tratemos el error de muestreo puesto que éste condiciona, y en gran parte determina, el tamaño de la muestra.

4.1. Métodos de muestreo

Anteriormente hemos señalado diferentes métodos de muestreo aleatorio, que ahora pasamos detenidamente a estudiar.

4.1.1. Muestreo aleatorio simple

La selección aleatoria de los sujetos que van a formar la muestra es el procedimiento más utilizado para conseguir que ésta sea representativa de la población de la que fue extraída.

Los requisitos exigidos para poder elegir muestras aleatorias son: primero, que todos los sujetos pertenecientes a la población tengan la misma probabilidad de ser obtenidos para formar la muestra; y segundo, que la selección de un sujeto no influya de ninguna manera sobre la selección de otro. Sólo si se han conseguido estos dos requisitos podemos asegurar la aleatoriedad de la muestra.

El proceso a seguir en la aplicación de esta técnica es:

1.º Definir la población.

Si es posible registrar todos los elementos, el procedimiento a seguir para seleccionar la muestra puede ser, o cualquier método mecánico, tipo lotería o similares, o la tabla de números aleatorios.

Si la definición de la población impide registrar todos los elementos que la constituyen, pero si conocemos el tipo de distribución que tiene la característica estudiada, habrá que comprobar si dicha característica, en la muestra obtenida, se distribuye de la misma manera que en la población de la que fue extraída.

2.º Enumerar a todos los individuos que forman la muestra, asignando a cada miembro de la población un número de identificación.

3.º Selección de la muestra.

Una vez enumerados todos los elementos, se seleccionan al azar los números de identificación de los sujetos, que pasarán a formar la muestra.

Ejemplo de selección de una muestra a través de la tabla de números aleatorios: En un censo de 150 alumnos se desea realizar una investigación, para la que se necesita seleccionar una muestra de 50 sujetos.

El proceso a seguir es el siguiente:

Se comienza por asignar números a los 150 alumnos desde el 001, 002, 003... etc., hasta el 150. A continuación se utiliza la tabla de números aleatorios, que son listas de números al azar, distribuidos en columnas y filas (Buendía 1994a: 91-93). Se determina, aleatoriamente, el punto por donde se va a comenzar a leer en las tablas y se anotan los números que progresivamente van apareciendo, hasta completar los 150 que nos interesan para la muestra. En la lectura y correspondiente anotación de los números de tres cifras se eliminarán aquellos que salgan repetidos, y los que tengan un valor superior a 150; número máximo de la población definida.

Cabe esperar que la muestra así extraída sea representativa.

Tiene el inconveniente de lo costoso y difícil de utilizar para poblaciones grandes, por la dificultad de realizar un listado de todos los elementos y por lo tedioso que resultaría la extracción de una muestra. Por eso, en este caso, se recurren a otros procedimientos de muestreo.

4.1.2. Muestreo aleatorio sistemático

Es, realmente, una variante del muestreo aleatorio simple, pero las unidades muestrales se seleccionan a partir del coeficiente de elevación y se elige, como punto de partida para proceder a la selección, un número no superior a dicho coeficiente. Los números que forman la muestra serán los resultantes de sumar el coeficiente de variación progresivamente, a partir del elegido aleatoriamente.

Coeficiente de elevación: N/n .

N = número de unidades del universo.

n = número de elementos de la muestra.

Ejemplo: Se desea seleccionar 2.000 alumnos de Educación Secundaria de una población de 24.000.

Se comienza calculando el coeficiente de elevación: $N/n = 12$.

Al número aleatoriamente elegido se le suma sucesivamente 12.

Si el número elegido al azar fue el 9, la serie que constituye la muestra sería: 9, 9 + 12; 21 + 12; 33 + 12, y así sucesivamente.

4.1.3. Muestreo aleatorio estratificado

Cuando la característica que nos interesa estudiar no se distribuye homogéneamente en la población, por la existencia de grupos o estratos entre sí heterogéneos, se efectúa un muestreo estratificado. El resultado es una muestra compuesta por tantas muestras, elegidas al azar, como estratos existan en la población. Los alumnos de un centro los podemos agrupar por edades, clase social, lugar de procedencia, etc. (estratos). Cada uno de estos grupos se pueden subdividir en subgrupos y elegir la muestra en función de ellos, si nos interesa una representación proporcional, de cada uno de ellos, en la muestra total.

El número de elementos del estrato podrá ser diferente desde el momento que es diferente el número total de sujetos que forman cada uno de ellos.

Para realizar este tipo de muestreo adecuadamente es necesario conocer el concepto de *afijación de una muestra*.

Por afijación se entiende la distribución de las encuestas entre los diferentes subconjuntos en que puede dividirse la población. Hay tres criterios de realizar la afijación: simple, proporcional y óptimo.

La afijación simple consiste en realizar el mismo número de entrevistas en cada subconjunto o estrato. En la afijación proporcional, las entrevistas se realizan proporcionalmente al peso de la población en cada estrato. En la afijación óptima, se contempla la homogeneidad o heterogeneidad de la población en una determinada característica. La mayor heterogeneidad en la categoría exige mayor número de casos.

Cuando se utiliza la técnica de afijación no proporcional, las fracciones de muestreo son distintas y esto no afecta si se tabulan por separado cada submuestra; pero si la tabulación se realiza conjuntamente, habrá que proceder a una ponderación de las diferentes submuestras si no se quieren deformar las estimaciones.

Ejemplo

Tipo de afijaciones en una muestra de profesores de universidad (la desviación típica se hace en función de los años de docencia).

Estratos	Porcentaje profesores	Desviación típica	Afijación		
			1	2	3
1. Catedráticos	10	5	250	100	350
2. Titulares	50	7	250	500	500
3. Asociados	30	3	250	300	100
4. Ayudantes	10	1	250	100	50
Total	100		2.000	2.000	2.000

Nota: Los datos no son reales y se han utilizado sólo a título de ejemplo.

La ponderación consiste en devolver a cada estrato la proporcionalidad que realmente tiene respecto a la población de la cual fue extraída. La ponderación se obtiene dividiendo el porcentaje que representa la población de cada submuestra sobre el universo total, por el porcentaje que representa en la muestra.

Ejemplo de cálculo del coeficiente de ponderación (Rodríguez, 1991: 62): Se recogen los datos básicos de un estudio electoral realizado en una comunidad autónoma con dos provincias, de las cuales, en el ejemplo, una tiene 500.000 electores y la otra 1.000.000. Dado que se necesitan estimaciones a nivel provincial, se ha diseñado una muestra de 1.000 entrevistas, dividida en dos submuestras de 500 entrevistas. Se organizaron como aparece en la siguiente tabla:

Provincias				
	A	B	Total	
1. Electores				
	% sobre T (1.500)		500.000	1.000.000
			33,3	66,6
2. Muestra				
	% sobre TM (1.000)		500	500
			50	50
3. Votantes del partido X en la muestra				
	% sobre M (500)		300	100
			60	20
4. Ponderación cociente entre % sobre T y % sobre TM				
			0,66	1,332

Nota: T: elementos totales del universo; TM: elementos totales de la muestra; M: muestra.

4.1.4. Muestreo aleatorio por conglomerados

Cuando la selección no se realiza sobre individuos independientes sino sobre grupos de individuos que presentan alguna afinidad, siendo este conjunto la unidad de selección, se realiza un muestreo por conglomerados.

Las unidades de selección pueden ser naturales, por ejemplo, los departamentos universitarios, los centros de enseñanza primaria o los hospitales; y artificiales, por ejemplo, las urnas electorales.

Definidos los conglomerados, la forma de extraer la muestra depende del tamaño del conglomerado. Si el número de elementos es pequeño, se pueden realizar entrevistas a todos los elementos del conglomerado, pero si es numeroso, habrá que proceder a seleccionar submuestras en cada uno.

Por ejemplo, si se desea conocer los hábitos higiénicos de los alumnos de Educación Secundaria Obligatoria de los centros escolares de Andalucía, se podría realizar el siguiente muestreo:

- *Universo*: Las unidades —clases— de primero, segundo y tercero de BUP de los centros escolares del municipio de Madrid.
- *Estratificación*: Los centros se estratificaron según su tamaño y su pertenencia al sector público o privado. Del cruce de estas dos variables se obtuvieron los nueve estratos que aparecen en el cuadro.
- *Tamaño de la muestra*: Se fijaron 266 unidades, distribuidas de forma proporcional al peso de cada estrato.
- *Selección de los elementos de la muestra*: A partir de un listado de las clases por centros y estratos, se seleccionaron las unidades que deberían formar parte de la muestra, mediante un muestreo aleatorio simple. Los cuestionarios se aplicaron a todos los alumnos de las clases seleccionadas.

Tamaño	Tipo de centro			Total
	Público	Privado religioso	Privado no religioso	
Pequeño	14	308	318	640
Mediano	180	188	240	608
Grande	700	18	34	752
Total	894	514	592	2.000

Este muestreo tiene el inconveniente de una importante pérdida de información, al no ser los individuos las unidades muestrales sino el conglomerado. Tiene la ventaja de la simplificación del proceso de extracción de la muestra. Sin embargo, cuando los elementos de los conglomerados son muy numerosos, hay grandes inconvenientes para aplicar el muestreo de conglomerados de una sola etapa, dado que el

número de entrevistas sería muy numeroso y las estimaciones, por otra parte, serían poco precisas. En estos casos se opta por una variedad del muestreo aleatorio por conglomerados que es el muestreo polietápico.

Ejemplo: si queremos realizar un trabajo de encuesta con profesores universitarios, éstos están adscritos a Universidades que a la vez se dividen en Facultades y éstas a su vez en Departamentos, y a ellos pertenecen los profesores. Con este proceso hemos elegido entre diferentes conglomerados y a la vez en diferentes etapas. Las unidades de la primera etapa se llaman unidades primarias de muestreo (ejemplo: Universidades); a las unidades subsecuentes, unidades de segundo, tercero, etc., orden (ejemplo: Facultades), y la que forma la muestra de estudio se denomina unidad última de muestreo; en este caso, el Departamento. La selección en cada etapa se realiza de forma aleatoria. Los individuos de los departamentos seleccionados constituirán la muestra de estudio.

4.2. Encuestas no representativas

Cuando el muestreo no es aleatorio, los resultados obtenidos con el método de encuesta podrán describir o explicar lo ocurrido en una situación determinada, pero no podrán extrapolarse a toda la población, dado que no se puede determinar la probabilidad de que todos los elementos de la población hayan estado disponibles para ser seleccionados.

Este muestreo se denomina incidental o accidental, y el investigador lo realiza aprovechando los elementos de la población que le son fácilmente accesibles. Este muestreo obedece más a la facilidad de realización que a una cuidadosa planificación para conseguir la representatividad. Las muestras obtenidas por este sistema no son aleatorias y por tanto los resultados obtenidos con ellas no pueden ser generalizables a ninguna población.

4.3. Tamaño de la muestra

El tamaño que debe tener la muestra que se selecciona depende básicamente del tipo de estudio que se vaya a realizar.

En los trabajos experimentales la muestra puede ser bastante más pequeña que si realizamos trabajos descriptivos, en los que la única forma de control es aumentando el número de elementos en la muestra.

Desde la perspectiva estadística, el tamaño de la muestra está directamente asociado al grado de precisión que se desea obtener en la estimación de los parámetros de la población. La precisión depende de la cantidad máxima de error que estamos dispuestos a tolerar en las estimaciones muestrales para un determinado intervalo de confianza. En función de esto, el tamaño de la muestra va a estar condicionado por la cantidad máxima de error que se admita en el verdadero valor de la población, y en segundo lugar, de los límites o intervalos de confianza que se han fijado para la estimación de dicho valor.

Para estimar el tamaño de la muestra de una población, vamos a comenzar diferenciando universos pequeños, equivalentes en la terminología estadística a poblaciones finitas, y universos grandes, equivalentes a poblaciones infinitas.

4.3.1. Tamaño de la muestra en poblaciones finitas

Para universos pequeños la fórmula para estimar la media es:

$$n = \frac{N^2 K^2 \sigma^2}{Ne^2 + K^2 \sigma^2}$$

n = Tamaño de la muestra.

N = Tamaño del universo.

K = Nivel de confianza.

σ^2 = Cuasivarianza de la población.

e = Error de muestreo.

Como la σ^2 de la población no se conoce, se suele estimar a partir de la cuasivarianza de la muestra.

Fórmula:

$$S^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

Ejemplo: Deseamos calcular el tamaño de la muestra para estimar el número de horas diarias que ven la televisión los alumnos de Educación Primaria de la capital A.

Datos: Número de alumnos = 1.200; σ^2 de la población = 1,8 (se puede obtener a través de un sondeo piloto y estimarla a partir de esos datos); $K = 2 = 0,95$ de probabilidad; error de muestreo (prefijado) = 0,10.

$$n = \frac{1.200 \times 2^2 \times 1,8}{1.200 \times 0,10^2 + 2^2 \times 1,8} = 45$$

La muestra ha de tener 45 elementos si admitimos un error de muestreo de 0,10. La fórmula para conocer el error de muestreo en la estimación de medias es:

$$e_{\bar{x}} = \sqrt{\frac{N - n}{N} \cdot \frac{\sigma^2}{n}}$$

Las fórmulas son diferentes si lo que queremos es conocer el tamaño de la muestra para estimar proporciones. Una explicación detallada de estos casos puede verse en Rodríguez (1991: 65-68), Buendía (1994a: 94-95).

4.3.2. Tamaño de la muestra en poblaciones infinitas

Cuando tratamos con universos grandes, si la fracción de muestreo no supera el 5 por 100, la fórmula para calcular el tamaño de la muestra se simplifica considerablemente. Exponemos a continuación un supuesto, para aplicar en la práctica, cuando se desean estimar proporciones. La fórmula es:

$$n = \frac{K^2 P(1 - P)}{e^2}$$

n = Tamaño de la muestra.

P = Proporción de una categoría o probabilidad de ocurrencia de esa categoría.

$P(1 - P)$ = Proporción de la categoría contraria.

K = Nivel de confianza.

e = Error de muestreo.

En este caso $P(1 - P)$ es la varianza, por tratarse de variables que se distribuyen binomialmente.

Para calcular el error de muestreo de la proporción, la fórmula es:

$$e_p = \sqrt{\frac{P(1 - P)}{n}}$$

e_p = Error de la estimación de la proporción.

n = Tamaño de la muestra.

P = Proporción de una categoría de la variable.

$P(1 - P)$ = Varianza.

Ejemplo: Se desea calcular el tamaño de la muestra para estimar proporciones cuando la probabilidad de ocurrencia de la categoría es 0,30 y el error de muestreo prefijado es 0,02. El nivel de confianza es del 95 por 100, o sea, $K = 2$.

$$n = \frac{2^2 \times 0,30 \times 0,70}{0,02^2} = 210 \text{ elementos}$$

No obstante, existen tablas para calcular el tamaño de la muestra para universos grandes, con un nivel de confianza del 95,5 por 100, y en el supuesto de muestreo aleatorio simple, véase Rodríguez Osuna, 1991: 72.

5. OPTIMIZACIÓN DEL PROCESO

En la investigación por encuesta, la calidad del proceso y consecuentemente de los resultados puede verse amenazada si no se toman las precauciones oportunas tanto por parte del encuestador como en la confección del instrumento y selección de la muestra.

Si los problemas vienen ocasionados por el proceso seguido en la selección de la muestra, los resultados no podrán ser generalizados a ninguna población. Los errores derivados de una inadecuada selección muestral no invalidan el instrumento, pero sí los resultados obtenidos con referencia a determinados parámetros.

El mayor error debido al muestreo es el error muestral. Éste consiste en el error que se comete cuando se trabaja con muestras aleatorias, por el mero hecho de no trabajar con toda la población. Por ejemplo, si elegimos tres o cuatro muestras de una población, aunque todas hayan sido elegidas correctamente, difícilmente coincidirán entre ellas, ni totalmente con la población. Estas variaciones se producen accidentalmente en los muestreos aleatorios y hace que las estimaciones sean más imprecisas que si no existiera este error.

Los errores ajenos al muestreo pueden ser más numerosos y de mayor efecto que los muestrales (Gómez, 1990); están constituidos por problemas no resueltos que se presentan en todas las etapas de la encuesta y plantean graves problemas de fiabilidad y sobre todo de validez (Kish, 1972).

La fiabilidad puede ser entendida como la credibilidad de los resultados tanto por la correcta interpretación del encuestado como por la ausencia de error en la interpretación de los encuestadores.

Cuando la encuesta incluye cuestiones delicadas, que pueden inducir a problemas de posible infiabilidad, se recomienda realizar varias preguntas alternativas, de similar contenido, que aseguren la fiabilidad de la respuesta. En cuanto al encuestador, es preciso una buena formación que evite los sesgos propios de la interpretación de las respuestas.

La falta de validez de los resultados puede ser debida tanto a un mal diseño de la encuesta como a problemas de selección muestral o características del entrevistador.

Durante el diseño de la encuesta es importante cuidar la presentación de las cuestiones, para que realmente recojan el objetivo que pretendemos cubrir con la investigación. La facilidad de comprensión de cada una de las cuestiones, la longitud de las preguntas y respuestas, la forma de presentarlas, etc., pueden impedir que los resultados sean válidos.

En la entrevista, los dos interlocutores pueden influirse mutuamente. Las características personales de ambos pueden ser decisivas para las informaciones recogidas. La manera de preguntar, la actitud del entrevistador, el clima que se genera, el tipo de preguntas, etc., determinan en gran medida las respuestas y la fiabilidad de las mismas. En determinadas situaciones se impone una cuidadosa selección para evitar reacciones adversas en los entrevistados, o respuestas en función de la actitud expectante del entrevistador.

6. CODIFICACIÓN Y ANÁLISIS DE LOS DATOS

El paso previo al análisis de datos de cualquier encuesta es la codificación. Por ésta se entiende la asignación de números tanto de identificación de los sujetos como a las diferentes preguntas y respuestas del cuestionario. Las preguntas con más de una respuesta exigen la consideración de cada respuesta como una variable distinta y por tanto se realizará diferente codificación con cada una de ellas.

Esta etapa de preparación de los datos es sumamente importante por los errores que se pueden introducir en la investigación, debido a una mala preparación de los codificadores o a omisiones involuntarias que podrían falsear todo el proceso.

Ejemplo de codificación:

A) Pregunta de una sola respuesta					
Sujetos					
Cuestión 4: Tipo de centro al que pertenece		(1)	(2)	(3)	(4)
Público	(1)	X	X		
Privado	(2)			X	
Concertado	(3)				X
B) Pregunta con diferentes respuestas					
Sujetos					
Cuestión 6: Técnicas y procedimientos de evaluación		(1)	(2)	(3)	(4)
Observación	(1)	X		X	
Pruebas objetivas	(2)		X	X	X
Escalas de estimación	(3)	X			
Cuadernos de clase	(4)	X		X	X
Exámenes orales	(5)		X		
Autoevaluaciones	(6)				
Informes de los padres	(7)	X	X		

Los datos recogidos en estos dos ítems (el 4 de respuesta única y el 6 con siete opciones: a, b, c, d, e, f, g) se podrían organizar como sigue:

Sujeto	P01	P06a	P06b	P06c	P06d	P06e	P06f	P06g
01	1	1		1	1			1
02	1		1			1		1
03	2	1	1		1			
04	3		1		1			

Una vez codificados y preparados los datos se pueden realizar diferentes tipos de análisis según interese un primer nivel de descripción o pasar a explicar las relaciones establecidas en las hipótesis.

La técnica estadística elegida para el análisis de datos dependerá de los objetivos del estudio. Si es describir una característica de la muestra en un momento determinado, se utiliza estadística univariante; si se pretende describir y/o explicar las relaciones de varias características al mismo tiempo, se recurre a la estadística bivariente o multivariante.

Hoy, gracias al software estadístico, el proceso es fácil y de gran rapidez. Entre los paquetes más utilizados están: SPSS, BMDP, SYSTAT, LISREL, etc.

Cuando los datos han sido recogidos con entrevistas, tendremos los pasos previos de transcripción y preparación de la información. Si las entrevistas son estructuradas, se puede proceder como en los cuestionarios, considerando las diferentes posibilidades de respuestas como diferentes variables, y utilizar para las descripciones y/o explicaciones los paquetes estadísticos anteriormente comentados.

Se puede también utilizar otro tipo de software como el NUDIST, ETNOGRAF, AQUAD, SPAD.N y SPAD:N Integre (Tesch, 1990; Etxeberria, García, Gil y Rodríguez, 1995; Miles y Huberman, 1984; Colás y Buendía, 1994, etc.), que aunque se denominan comúnmente paquetes para el análisis de datos cualitativos, son, realmente, análisis cuantitativos de textos escritos.

7. ELABORACIÓN DE CONCLUSIONES

El último paso del proceso de investigación consiste en la interpretación de los datos y la posterior elaboración de las conclusiones del estudio. Estas, en función de los objetivos de la investigación, pueden ser de dos tipos. El primero es una descripción o comportamiento de la/s variable/s objeto de estudio; y el segundo el establecimiento de relaciones entre variables en términos de concomitancia.

- Uno de los mayores errores de interpretación en los estudios por encuesta aparece en la generalización de resultados. Cuando las muestras no han sido adecuadamente seleccionadas, las deducciones teóricas sólo se podrán generalizar a otras muestras de características similares a las recogidas en el estudio. Así pues, al elaborar las conclusiones se debe tener en cuenta el tipo de muestra con la que se ha obtenido las conclusiones, las deducciones teóricas de la muestra seleccionada.
- Un análisis de las discrepancias, si las hubiera, entre las previsiones realizadas en las hipótesis y los resultados obtenidos.
 - Una comparación de los resultados obtenidos en la investigación y los obtenidos por otros investigadores. Si los resultados fueran discrepantes, deben analizarse las posibles fuentes de las inconsistencias.
 - Es importante realizar sugerencias para nuevas investigaciones y dejar abiertos nuevos interrogantes que permitan ampliar el trabajo realizado.

En la redacción de conclusiones se debe ser cauteloso y no inferir relaciones que no habrían sido contempladas en las hipótesis, o ir más allá de lo que los datos analizados nos permiten. A veces un deseo de verificación lleva al investigador a aceptar márgenes de error mayores de lo que es posible con el nivel de significación establecido; olvidando que el proceso de investigación está encaminado a la confirmación o no de las hipótesis, siendo igual de válidas si las confirma o las refuta.

En la interpretación de los datos y posterior elaboración de conclusiones pueden aparecer dudas que tendrán que ser planteadas como nuevos problemas de investigación. Un problema investigado no tiene por qué ser un problema concluido. La investigación debe generar nuevos interrogantes cuyo estudio permite ampliar el cuerpo de conocimientos científicos.

8. EL INFORME DE INVESTIGACIÓN

La función del informe de investigación es comunicar al resto de la comunidad científica el problema investigado, los resultados obtenidos, así como la presentación de nuevas técnicas y enfoques metodológicos. Esta información debe realizarse de la forma más clara, objetiva y concisa posible.

El éxito de la investigación viene condicionado por el conocimiento que de la misma tenga la sociedad y el cumplimiento de los fines para los que fue elaborada. Por esto, en el informe de investigación, tan importante son los resultados obtenidos como el desarrollo del proceso seguido para conseguirlos. Para una mayor claridad y concisión, sobre todo en los informes para la sociedad en general, a través de diarios u otros medios de comunicación debe incluirse una ficha técnica de la encuesta, antes de los comentarios derivados de la misma.

8.1. Organización del informe

En la organización del informe se siguen prácticamente los mismos pasos que en el proceso de investigación.

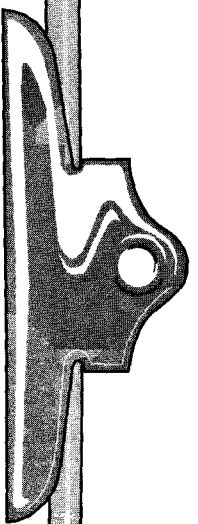
Básicamente los tres grandes apartados son:

- Planteamiento del problema.
- Metodología.
- Básicamente los tres grandes apartados son:

- Planteamiento del problema.

• Metodología. En la organización del informe se siguen prácticamente los mismos pasos que en el proceso de investigación. Básicamente los tres grandes apartados son: Planteamiento del problema. Metodología. Básicamente los tres grandes apartados son: Planteamiento del problema.

- Metodología. En la organización del informe se siguen prácticamente los mismos pasos que en el proceso de investigación. Básicamente los tres grandes apartados son: Planteamiento del problema. Metodología. Básicamente los tres grandes apartados son: Planteamiento del problema.



1. TÍTULO Y AUTOR	_____
2. RESUMEN	_____
3. INTRODUCCIÓN	_____
3.1. Planteamiento del problema	_____
3.2. Revisión de la literatura	_____
3.3. Hipótesis de la investigación	_____
4. MÉTODO	_____
4.1. Definición de la población y de la muestra:	_____
Método de muestreo y error muestral	_____
4.2. Instrumentos de recogida de datos	_____
4.3. Características técnicas de los instrumentos	_____
5. RESULTADOS	_____
5.1. Análisis de resultados	_____
5.2. Conclusiones	_____
6. REFERENCIAS BIBLIOGRÁFICAS	_____

Por ejemplo, el título: «Estudio sobre la relación del rendimiento académico y la clase social», puede ser sustituido por este otro: «Clase social y rendimiento académico». En ambos casos se ponen de manifiesto las dos variables objeto de estudio, pero, en el primer caso, se utilizan palabras que alargan el título innecesariamente.

A continuación del título debe aparecer el autor o autores, seguido del centro donde se ha llevado a cabo la investigación.

- **Resumen:** El resumen es un párrafo de 12 a 15 líneas, que contiene la información más importante del trabajo realizado. En él deben figurar las variables que se han estudiado, el procedimiento que se ha seguido, los principales resultados y las conclusiones.

Debe ser construido cuidadosamente porque de él va a depender que el lector se interese o no por la investigación.

- **Introducción:** El objetivo de este apartado es facilitar una precisa información sobre el marco teórico en el que se inserta el problema, además de una especificación del mismo y de las hipótesis que se pretenden validar.

Con el comentario de la literatura existente sobre el tema se sientan las bases teóricas del problema, y por otra parte se le proporciona al lector lo que ya está investigado y los interrogantes que quedan por resolver. En este apartado se podría seguir el siguiente orden (Buendía, 1994c):

1. Amplia exposición de la problemática.
2. Discusión del origen del problema y las concepciones teóricas más importantes.
3. Análisis de las teorías más relevantes.
4. Breve descripción de los trabajos realizados sobre la temática.
5. Presentación sucinta del estado actual del conocimiento sobre el problema.
6. Derivación del problema a partir de la presentación anterior.

- **Método:** La primera parte debe dedicarse a la muestra sobre la que se ha hecho el estudio, el método de muestreo y el nivel de significación establecido.

En la descripción se incluirá el número de elementos que la forman, las características comunes más importantes, el procedimiento de selección y asignación de los sujetos y el nivel de significación establecido.

A veces en el informe se incluye una ficha técnica como la que exponemos a continuación:

Ejemplo de ficha técnica (Juste, Ramírez y Barbadillo, 1991; citados por León y Montero, 1993: 79).

Ficha técnica

Características de la muestra

- **Ámbito:** Nacional. Se incluyen las provincias insulares y se excluyen Ceuta y Melilla.

- **Universo:** Población española de ambos sexos de 18 años y más.
 - **Tamaño:** 2.500 entrevistas.
 - **Afijación:** Proporcional.
 - **Puntos de muestreo:** 155 municipios y 45 provincias.
 - **Procedimiento de muestreo:** Polietápico, estratificado por conglomerados, con selección de las unidades primarias de muestreo (municipios) de forma aleatoria proporcional, de las unidades secundarias (secciones) de forma aleatoria simple y de las unidades últimas (individuos) por rutas aleatorias y cuotas de sexo y edad.
- Los estratos se han formado por el cruce de las 17 comunidades autónomas con el tamaño del hábitat, dividido éste en siete categorías: menos de 2.000 habitantes; de 2.001 a 10.000; de 10.001 a 50.000; de 50.001 a 100.000; de 100.001 a 400.000; de 400.001 a 1.000.000; más de 1.000.000 de habitantes.
- **Error muestral:** Para un nivel de confianza del 95,5 por 100 (dos sigmas) y $P = Q$ el error es ± 2 para el conjunto de la muestra.
 - **Fecha de realización:** Días 16-20 de junio de 1990.

Instrumentos de recogida de datos

Deben describirse detenidamente los instrumentos de obtención de datos; indicando: el diseño del cuestionario, el contenido y el tipo de preguntas, las características técnicas del instrumento y las modificaciones realizadas, si hubo encuesta piloto.

Resultados

Los resultados deben exponerse de la forma más completa y precisa posible. Es conveniente sintetizar los datos en tablas, con los resultados obtenidos del análisis estadístico. La exposición de estos datos se realizará ordenadamente de manera que vayan respondiendo a las hipótesis planteadas.

Debe indicarse el nombre de la prueba estadística utilizada, el nivel de confianza, los grados de libertad, etc. Las hipótesis se irán aceptando o rechazando progresivamente en la medida que se analizan los resultados obtenidos. Si aparece algún resultado interesante, que incluso abre nuevos interrogantes en la investigación científica, deberá consignarse como posibles problemas de investigación o nuevas líneas de trabajo.

Referencias bibliográficas

El último apartado del informe se dedica a las referencias bibliográficas. Éstas son listas de los autores citados a lo largo de todo el trabajo.

Las listas deben realizarse por riguroso orden alfabético del primer apellido de los autores, tanto si son de libros como de artículos. Estas citas deben contener todos los datos necesarios para poderlas identificar en cualquier momento.

Ejemplos:

- a) **Cita de un libro.** Autor (año de publicación). Título con letra cursiva. Lugar de edición: editorial.

Si hay más de un autor deben indicarse todos separados por comas, excepto el último que va precedido de la conjunción «y».

Colas, P., y Buendía, L. (1992). *Investigación Educativa*. Sevilla: Alfar.

Si no se trabaja con la edición original, debe ponerse la edición correspondiente después del título.

Colas, P., y Buendía, L. (1994). *Investigación Educativa*. Segunda edición. Sevilla: Alfar.

Cuando se trata de un libro traducido, la cita seguirá el mismo orden, añadiendo al final entre paréntesis la fecha de la edición original.

Weis, C. H. (1975). *Investigación Evaluativa*. México: Trillas (Versión original, 1972).

En los libros traducidos también se puede realizar la cita con el título original y al final entre paréntesis la editorial y el año de la traducción.

Cuando son compilaciones, después del nombre del autor, entre paréntesis, se especificará la función.

Buendía, L. (Coord.) (1993). *Análisis de la investigación educativa*. Granada: Servicio de Publicaciones de la Universidad de Granada.

Cuando la referencia es del capítulo de un libro, se cita el autor del capítulo, fecha y título; a continuación el nombre del director o coordinador, título del libro, páginas del capítulo de referencia, lugar de edición y editorial.

Bartolomé, M. (1978). Estudio de las variables en la investigación en educación. En J. Arnan (Dir.). *Métodos de investigación en las ciencias humanas* (pp. 103-138). Barcelona: Omega.

- b) **Cita de un artículo.** En este caso lo que va en letra cursiva es el título de la revista. Se debe especificar el volumen de la revista y las páginas que ocupa el artículo, separando la primera de la última mediante un guión.

Jacob, E. (1987). Qualitative research traditions: A review. *Review of Educational Research*, 57 (1), 1-50.

- c) **Cita de una ponencia o comunicación presentada en un congreso.** Se especificará el autor, año, título de la ponencia, nombre del congreso y a ser posible mes de celebración.

De Miguel, M. (1987). *Paradigmas de la investigación educativa*. II Congreso Mundial Vasco. Octubre.

En algunos informes después de este apartado se incluyen los anexos. Pueden incluir: el cuestionario completo, cuadros estadísticos y gráficas que no se incluyeron en los resultados, así como otros materiales de interés para el lector.

9. BIBLIOGRAFÍA

- Anguera, M. T. (1990). Metodología observacional. En J. Arnaú, M. T. Anguera y J. Gómez. *Metodología de la Investigación en Ciencias del Comportamiento*. Murcia: Universidad de Murcia.
- Buendía, L. (1991). Actitud de los Profesores ante la Reforma del Sistema Educativo. *Actas de las II Jornadas sobre la Reforma del Sistema Educativo*. Granada: ICE.
- Buendía, L. (1994a). El Proceso de Investigación. En P. Colás y L. Buendía. *Investigación Educativa* (pp. 69-110). Sevilla: Alfar.
- Buendía, L. (1994b). Técnicas e instrumentos de recogida de datos. En P. Colás y L. Buendía. *Investigación educativa* (pp. 201-244). Sevilla: Alfar.
- Buendía, L. (1994c). *Análisis de la Investigación Educativa*. Granada: Servicio de Publicaciones de la Universidad.
- De la Orden, A. (1988). Informática e Investigación educativa. En I. Dendaluce (Coord.). *Aspectos metodológicos de la Investigación Educativa* (pp. 276-295). Madrid: Narcea
- Denzin, N. K., y Lincoln, Y. S. (1994). *Handbook of Qualitative Research*. Thousand Oaks, California: Sage.
- Dillman, D. A. (1983). Mail and other self-administered questionnaires. En P. H. Rossi, J. D. Wright y A. B. Anderson (Eds.). *Handbook of Survey Research* (pp. 359-377). New York: Academic Press.
- Dwyer, J. H. (1983). *Statistical models for the social and behavioral sciences*. New York: Oxford University Press.
- Etxeberria, J.; García, E.; Gil, J., y Rodríguez, G. (1995). *Análisis de datos y textos*. Madrid: Rana.
- Fowler, F. J. (1988). *Survey Research Methods*. Beverly Hills, California: Sage.
- Gómez, J. (1990). Metodología de encuesta por muestreo. En J. Arnaú, M. T. Anguera y J. Gómez. *Metodología de la Investigación en Ciencias del Comportamiento*. Murcia: Universidad de Murcia.
- Groves, R. M. (1979). Actors and questions in telephone and personal interview surveys. *Public Opinion Quarterly*, 43, 190-205.
- Haberlein, T. A., y Baumgartner, R. M. (1978). Factors Affecting Response Rates to mailed Questionnaires: Quantitative Analysis of the Published Literature. *American Sociological Review*, 43 (4), 447-462.
- Hyman, H. H. (1972). *Secondary Analysis of Sample Surveys: Principles, Procedures and Potentialities*. New York: John Wiley and Sons.
- Kiecolt, K. L., y Nathan, L. E. (1985). *Secondary Analysis of Survey Data*. Beverly Hills, California: Sage.
- Kish, L. (1972). *Muestreo de encuestas*. México: Trillas.
- Manheim, H. L. (1982). *Investigación Sociológica. Filosofía y Métodos*. Barcelona: CEAC.
- Miles, M. B., y Huberman, A. M. (1984). *Qualitative data analysis: A Sourcebook of New Methods*. Beverly Hills: Sage.
- Pérez Juste, R. (1984). *Pedagogía Experimental*. Madrid: UNED.
- Rodríguez, J. (1991). *Métodos de muestreo*. Madrid: Centro de Investigaciones Sociológicas.
- Ruiz, J. I. (1996). *Metodología de la Investigación cualitativa*. Universidad de Deusto: Bilbao.
- Schae, K. V. (1965). A general model for the study of developmental problems. *Psychological Bulletin*, 64, 92-107.
- Seltiz, C.; Wrightsman, L. S., y Cook, S. W. (1980). *Métodos de investigación en la relaciones sociales*. Madrid: Riaap.
- Shure, G. H., y Meeker, R. H. (1978). A minicomputer system for multiperson computer assisted telephone interviewing. *Behavior Research Methods and Instrumentation*, 10, 196-202.

- Sierra Bravo, R. (1985). *Técnicas de investigación social. Teoría y Ejercicios*. Madrid: Paraninfo.
- Simon, J. L. (1978). *Basic Research Methods in Social Sciences*. New York: Random House.
- Tejedor, F. J. (1984). *Análisis de varianza aplicado a la Investigación en Pedagogía y Psicología*. Madrid: Anaya
- Tesch, R. (1990). *Qualitative Research: Analysis, Types and Software Tools*. London: The Falmer Press.

LA INVESTIGACIÓN OBSERVACIONAL

LEONOR BUENDÍA EISMAN
Universidad de Granada

OBJETIVOS

1. Conocer el proceso de la metodología observacional, diferenciándola de otras metodologías científicas.
2. Elaborar diferentes tipos de registros observacionales.
3. Optimizar los registros, analizarlos e integrarlos en el marco teórico de referencia.
4. Elaborar un informe de investigación observacional.

CONTENIDOS

1. Introducción.
2. Observación participante y observación no participante.
3. El método de observación: Consideraciones previas.
4. Diseños en la investigación observacional.
5. El control de los sesgos en la observación.
6. Registro de las observaciones.
7. Análisis de los datos obtenidos mediante observación.
8. El informe de la investigación observacional.
9. Bibliografía

BIBLIOGRAFÍA BÁSICA

- Anguera, M. T. (1990). *Metodología observacional*. En J. Arnau, M. T. Anguera y J. Gómez. *Metodología de la Investigación en Ciencias del Comportamiento*. Murcia. Universidad de Murcia.
- Anguera, M. T. (Ed.) (1991). *Metodología observacional en la investigación psicológica*. Barcelona: PPU
- Bakeman, R., y Gottman, J. M. (1989). *Observación de la interacción: Introducción al análisis de secuencias*. Madrid: Morata.
- Buendía, L. (1994). Técnicas e instrumentos de recogida de datos. En P. Colás y L. Buendía. *Investigación Educativa* (2.ª ed.). Sevilla: Alfar.

1. INTRODUCCIÓN

Todos los hechos de la vida pueden ser objeto de observación, pero no todas las observaciones que se realizan son válidas para construir un conocimiento científico. Sólo la observación rigurosa y sistemática nos ayuda a avanzar en el conocimiento, describiendo situaciones, contrastando hipótesis y haciendo así aumentar el cuerpo de conocimientos científicos.

Es a partir de la década de los años sesenta cuando más proliferan los estudios que utilizan la observación como método para responder a problemas educativos y representar las interacciones e interactividades entre profesores y alumnos (Buedia, 1993).

Bassedas, Coll y otros (1984: 20) manifiestan en este sentido que «*la observación perspicaz rigurosa y sistemática es el instrumento indispensable para comprender el comportamiento del alumno en el transcurso de las tareas de aprendizaje y para modificar su contenido y presentación en consecuencia. La observación adquiere casi papel preponderante y los problemas de qué observar y cómo observar se convierten en las cuestiones esenciales de la evaluación formativa en el Parvulario y en el Ciclo Inicial*».

La observación, como técnica de recogida de datos, es utilizada frecuentemente tanto en la investigación experimental o cuasi experimental como en la etnográfica, sin embargo en este capítulo nos vamos a centrar en el proceso que se sigue con la observación como método científico. Se retoman los trabajos realizados (Weick, 1968; Anguera, 1982, 1988, 1990; Anguera (Ed.), 1991; Irwin y Bushell, 1984; Altmann, 1974; Coolican, 1990; Fassnacht, 1982; Buendía y Pegalajar, 1990a, 1990b; Bakeman y Gottman, 1989; Suen y Ary, 1989; Bakeman y Quera, 1997; etc.) y se incluyen los últimos avances tanto en la obtención de informaciones como en el análisis de las mismas, para que los alumnos puedan tener una visión completa del proceso que debe seguirse cuando se utiliza el método de observación para responder a un problema educativo de manera científica.

2. OBSERVACIÓN PARTICIPANTE Y OBSERVACIÓN NO PARTICIPANTE

Se pueden distinguir dos clases de observaciones. En una, la persona observa simplemente lo que sucede, sin limitar sus observaciones a detalles particulares del comportamiento previamente categorizados. En el segundo caso, un plan previamente preparado concentra la atención en ciertos aspectos de la conducta, sin interacción entre el observador y el sujeto o grupo observado.

Si bien la primera, denominada observación participante por el nivel de implicación del observador en el grupo objeto de observación, es la técnica por excelencia de recogida de información en la investigación cualitativa, la observación no participante ha sido ampliamente utilizada en el modelo racionalista de investigación, tanto en estudios experimentales como correlacionales, como técnica de recogida de datos.

En la investigación cuantitativa, el fenómeno de estudio es primeramente reducido a un número de variables medibles y conductas observables. Estas variables son definidas operativa o conceptualmente antes que la observación sea realizada. Con la observación se obtienen datos sobre dichas variables para determinar las relaciones hipotetizadas entre ellas. A través de técnicas estadísticas se confirman o no dichas relaciones. Su principal ventaja es la objetividad y replicabilidad. El observador es independiente de la situación de observación y observadores diferentes podrían supuestamente llegar a los mismos resultados. Tiene, entre otros inconvenientes, la enorme limitación de su aplicación en los problemas educativos. Estos son problemas complejos, no desprovistos de valores, y con la técnica de observación cuantitativa las situaciones son excesivamente simplificadas.

Como contraste la investigación cualitativa pretende comprender la situación de estudio. El observador participa como un miembro del grupo social; se integra en el grupo, toma notas de las situaciones y posteriormente las recrea en el proceso del estudio. Fetterman (1984) establece las siguientes características de la investigación cualitativa:

- Es propia de la fenomenología. Desde este posicionamiento el observador intenta entender el fenómeno social desde la perspectiva de una persona. El objetivo no es pues un frío conocimiento científico, sino un entendimiento del fenómeno.
- El conocimiento es holístico. Todas las observaciones e interpretaciones están dirigidas a entender las relaciones de los elementos dentro de todo el sistema.
- El principio de contextualización requiere que todos los datos sean considerados solamente en el contexto en el que fueron obtenidos.

Tiene grandes ventajas al proporcionar una visión global y holística del fenómeno; y se describen las relaciones tal y como ocurren en los contextos en los que se observa.

Algunos de los inconvenientes son:

- Requiere una gran atención y observadores muy expertos.
- Se requieren observaciones continuadas, por largos espacios de tiempo.
- Se recoge gran cantidad de material escrito que dificulta su organización e interpretación.
- Por último, la validez de los datos puede verse afectada por las características personales del observador y su mejor o peor integración en el grupo observado.

No obstante, estas diferencias metodológicas no impiden que la observación, como técnica de recogida de informaciones, cualitativas o cuantitativas, pueden ser complementarias y no excluyentes. De hecho en investigaciones cuantitativas, el primer nivel de observación suele ser una observación exploratoria, sin hipótesis definidas y sin estructuración previa de la observación. Son descripciones al servicio de un modelo racionalista de investigación. Igualmente en la metodología cualitativa puede utilizarse la observación no participante para la obtención de datos o para complementar otras técnicas.

Además de esta función de recogida de informaciones, tanto para la metodología cualitativa como para la correlacional y experimental, la observación es un proceso riguroso de investigación que permite describir situaciones y/o contrastar hipótesis, siendo por tanto un método científico de indagar en una realidad. Como tal, vamos a estudiarlo en este capítulo.

3. EL MÉTODO DE OBSERVACIÓN: CONSIDERACIONES PREVIAS

Si la investigación se realiza en contextos naturales, con carácter fundamentalmente ideográfico, la observación permitirá otra forma de investigar, que si bien carece de la validez interna de los diseños experimentales para poder establecer relaciones causales, gana en validez externa y posibilidades de estudio en el contexto que se generen. Utilizada con esta finalidad y cumpliendo determinados requisitos, pierde el carácter exclusivamente instrumental para ser un procedimiento científico de investigación.

Lo que caracteriza un método científico es «el modo como opera para alcanzar algún objetivo determinado» (Bunge, 1979: 22).

En la metodología observacional, las fases de ese modo, como en cualquier procedimiento científico, son:

- Delimitación del problema.
- Recogida de datos y optimización.
- Análisis de datos.
- Interpretación de resultados.

No obstante, el método de observación sistemática tiene importantes notas diferenciadoras, que no contribuyen en detrimento de su cientificidad, sino que habría que considerarlas como diferentes directrices dentro del proceso de investigación.

Desde esta perspectiva, la metodología observacional la definimos como un procedimiento por el que se pretende captar el significado de una conducta surgida en un contexto natural, con ausencia total de manipulación, y que tras un registro riguroso de las manifestaciones de esa conducta y el análisis de los mismos podemos describirla, analizarla o explicarla en el contexto que se generó. El proceso que se sigue es el mostrado en la Figura 1.

3.1. El problema de investigación como delimitador del método

Cualquier trabajo de investigación, independientemente del modelo que se siga, exige acotar el campo de interés, hasta llegar a una concreción tal que lo que comenzó siendo un área problemática o una temática interesante se materialice en un problema concreto al que pretendemos dar respuesta con un método científico.

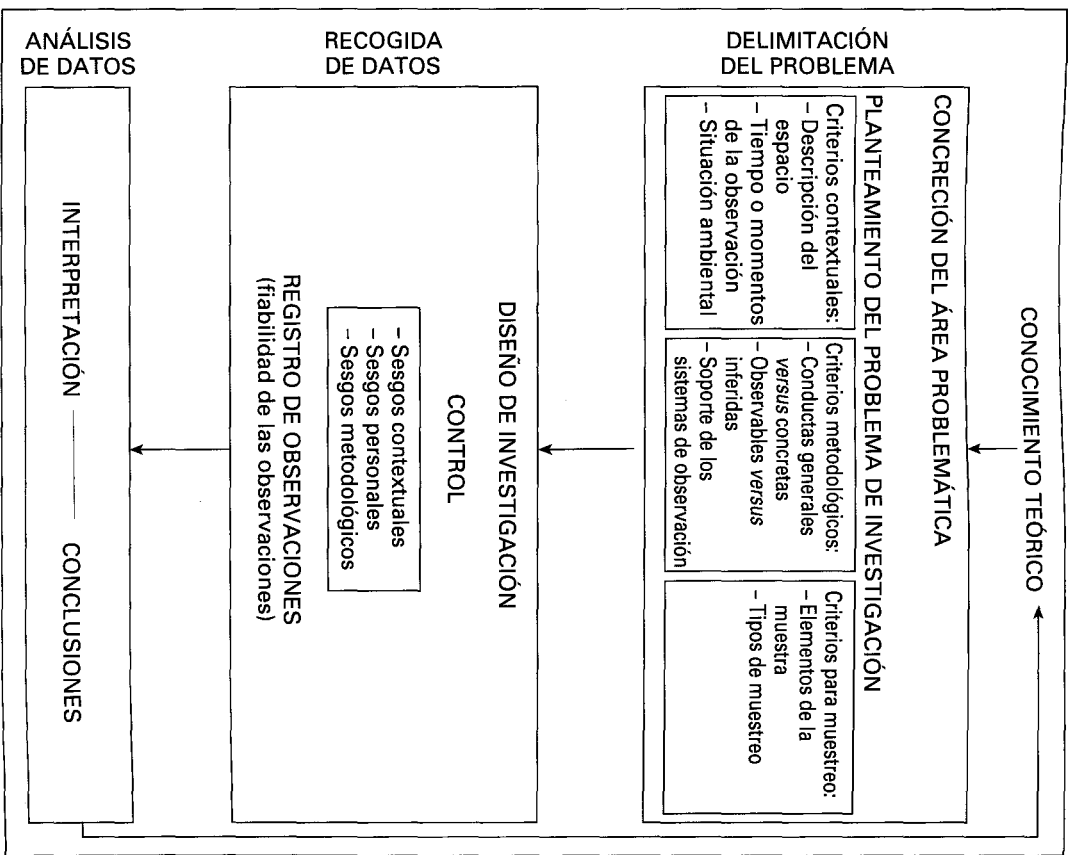


Figura 1

En la metodología observacional es frecuente llegar a esta concreción después de una observación general del área de interés. Esta observación preliminar la consideramos necesaria tanto si el problema surge de la propia experiencia del investigador en el contexto educativo, como si se genera por el conocimiento de teorías psicoeducativas que inducen al investigador a formular cuestiones que verificará con posterioridad.

La primera vía de generar problemas es la vía inductiva. En ella juega un importante papel la observación preliminar, porque será la fuente de preguntas y dudas que el investigador intentará resolver.

La segunda vía está basada en la deducción. A partir de una teoría previa, el investigador se hace preguntas, o intenta verificar las respuestas dadas en otras situaciones. En estos casos, la observación previa permite operarizar los términos del problema para la formulación posterior de las hipótesis, a la vez que se obtienen indicadores para organizar todo el proceso posterior.

Sobre estas conductas el acto de observación no puede ser un simple mirar y ver (Travers 1917: 146) preguntándose por la utilidad de esas observaciones. Esto que puede ser válido en la observación exploratoria, pierde su sentido una vez delimitado el objetivo de la investigación. Una observación científica necesita conjugar la percepción del hecho con la interpretación del mismo, teniendo en cuenta que éste no se da aislado, sino que forma parte de un marco definido que le da sentido a la interpretación realizada. En esta observación deberán realizarse los controles estadísticos necesarios para evitar los errores o distorsiones que se introduzcan en el acto de observar. Si cualquier actividad humana al ser percibida e interpretada puede presentar importantes variaciones, cuanto más no variará una actividad educativa en la que las conductas admiten múltiples interpretaciones por la diferente casuística de cada individuo y de cada momento, sin olvidar que el observador tiene su criterio personal y que a veces modifica, aunque sólo sea de forma mínima, lo percibido en dicho acto. Anguera (1988: 13) define el acto de observar como *la resultante de la actuación conjunta y necesaria de tres elementos: percepción, interpretación y conciencia previa, que darían lugar a una observación perfecta, incluso utópica, deteriorada en grado muy variable (y preferiblemente mínimo) al ser asumida por el observador humano, que genera unos sesgos o distorsiones sistemáticas o asistemáticas —según los casos— que habrá que tener en cuenta para incorporar adecuadamente mecanismos correctores.*

Para minimizar estos sesgos deberán tenerse en cuenta una serie de requisitos, que vamos a presentar agrupados en tres bloques:

- a) Criterios contextuales.
- b) Criterios metodológicos.
- c) Criterios para el muestreo observacional.

3.2. Criterios contextuales

Planteados los objetivos de investigación, y dada la importancia de la definición del contexto, el investigador tendrá que definir el lugar donde se va a realizar la observación y el momento o momentos (espacios temporales) de observación.

3.2.1. Descripción del espacio

El estudio y descripción detallada del lugar donde se va a realizar la observación es un requisito indispensable en la metodología observacional.

Uno de los problemas mayores en la metodología experimental es la reactividad: «distorsión del comportamiento de un sujeto debido a la alteración de la situación

natural en la que se desenvuelve habitualmente» (Riba, 1991: 43). En el método observacional, como sólo deben aplicarse los resultados al contexto concreto donde fueron obtenidos y a situaciones iguales o similares al lugar en el que se realizaron registros, esta reactividad es totalmente inexistente; no así otro tipo de reactividad que veremos más adelante (reactividad del observador y de los sujetos observados). Por esta razón, antes de realizar las observaciones es necesaria la definición del contexto para poder interpretarlas en el marco en el que fueron realizadas y extrapolarlas sólo a situaciones iguales o similares.

3.2.2. Momento o momentos de la observación

Establecer no sólo las unidades temporales o tiempo de la observación, sino además el criterio que se seguirá para conseguir sesiones lo más homogéneas posibles, es otro de los requisitos previos a la obtención de datos.

Si se desea observar el proceso que dos alumnos, trabajando conjuntamente, siguen en la resolución de un problema, además de definir el objeto del problema, el investigador establece las sesiones de observación y las variables que pueden influir y alterar la homogeneidad necesaria entre todas ellas. Si la sesión de observación anterior se realiza con el profesor «X», el cambio ocasional del profesor obliga a eliminar la sesión, por el interés que puede presentar para el investigador mantenerla constante. Igualmente debe ser eliminada la sesión que no es constante durante su desarrollo. Una interrupción de la secuencia durante más tiempo del previsto, en función del tema de estudio, aconseja eliminar la secuencia de observación y comenzar de nuevo.

3.3. Criterios metodológicos

Hemos agrupado en este apartado aquellos aspectos a definir o determinar y que la opción por uno u otro plantea cuestiones metodológicas que afectan al qué observar y cómo observar.

3.3.1. Conductas generales versus aspectos concretos

Por conducta general entendemos una manifestación amplia de conducta, capaz de ser dividida en partes para la observación, y que al ser interpretadas conjuntamente le confieren significado. Puesto que la percepción recae sobre los aspectos que la componen, el significado que se le atribuye exige un alto grado de inferencia. Es denominada por algunos autores conducta molar (Barker y Wright, 1955; Sackett, Ruppenthal y Gluck, 1978; Anguera, 1982; Suen y Ary, 1989).

Cuando se descompone una conducta global en diferentes dimensiones que son objeto de estudio y observamos esos aspectos realizando descripciones de los rasgos observados directamente, la inferencia que se realiza en la descripción es mínima o nula; dado que no es necesario realizar ningún tipo de abstracción para definirla, estamos observando rasgos o aspectos concretos, denominados por algunos autores conducta molecular (Anguera, 1982; Blanco y Anguera, 1991; Suen y Ary, 1989). La conducta general es ampliamente utilizada en estudios que exigen una visión global, de conjunto, del fenómeno observado.

Tiene el inconveniente de la subjetividad al no ser percibida directamente sino a través de las manifestaciones de las diferentes dimensiones, o rasgos, que la constituyen.

La observación de aspectos concretos, conducta molecular, es la unidad de observación por excelencia, al tener cada rasgo significado por el mismo, por lo que el nivel de abstracción es nulo. Sin embargo, al observar sólo un aspecto se pierde el sentido global, de conjunto, de gran importancia en los problemas educativos.

3.3.2. Conductas observables versus conductas inferidas

El objeto de observación puede ser una conducta observada directamente o bien conductas encubiertas cuyo conocimiento se inferirá de las manifestaciones indirectas de la conducta. Las conductas inferidas, al no ser observadas directamente, han sido objeto de numerosas críticas; pero temas importantes en educación, por ejemplo, las creencias de los profesores, los procesos de interacción e incluso el estado anímico, tendrían, bajo esta creencia, que ser olvidados por la metodología observacional.

Fassnacht (1982: 77-81), además de estas unidades de observación, especifica otras 17 unidades, algunas bipolares, más específicamente relacionadas con los sistemas de observación. Estas son:

- **Unidades naturales.** Son detectadas por la percepción y reflejadas en el lenguaje natural.
- **Unidades de conducta.** Una unidad de conducta puede ser descrita como una unidad natural.
- **Unidades inductivas versus deductivas.** Esta distinción se refiere a cómo son construidas las unidades. Inductivas, por vía empírica; deductivas, por vía racional a partir de una teoría.
- **Unidades observables directamente versus unidades inferidas.** Las unidades inferidas y no observadas directamente pueden ser inobservables por naturaleza, por ejemplo la emoción, o inobservable circunstancialmente, al no poder ser percibida en determinados momentos.
- **Unidades descriptivas versus evaluativas.** Esta distinción fue hecha primeramente por Boyd y DeVault (1960), según que las unidades observadas sean sólo para realizar descripciones o además interesen realizar juicios de valor.
- **Unidades fenomenológicas.** Se refiere a las conductas no verbales humanas.

- **Unidades morfológicas.** Son similares a las anteriores pero enfatizando aspectos formales o estructurales como criterio para seleccionar unidades de conducta.
- **Unidades extraídas del análisis factorial.** Son unidades extraídas de las dimensiones subyacentes en un análisis factorial.
- **Unidades discretas versus continuas.** Según se realicen conteos, sin valores intermedios, o mediciones continuas.
- **Unidades simples versus complejas.** La unidad compleja incluye diferentes niveles de información, a veces sólo perceptible por los observadores humanos. Ejemplo: Una sonrisa, según en qué contexto, puede expresar alegría, soma, desconfianza, etc.
- **Índices como unidades.** Tienen gran similitud con las unidades anteriores. El índice está compuesto de varios indicadores deducidos conjuntamente. Son utilizados fundamentalmente por economistas. El índice de coste de la vida es un ejemplo típico de estas unidades.
- **Unidades reduccionistas.** Son las mas pequeñas unidades de conducta con significado.
- **Unidades causales.** Las unidades de conducta con una causa común y que son consideradas idénticas.
- **Unidades funcionales.** Aunque este término tiene diferentes significados, todos tienen en común el énfasis sobre la importancia del contexto en la unidad de análisis.
- **La situación como unidad.** En Sociología y cada vez más en Educación, la situación como unidad de observación aparece en los estudios. Son unidades que comprenden la situación y la conducta (Barker y Wright, 1955), o bien sólo la situación.
- **Unidades molares.** Ya descritas anteriormente, son integradoras y tras diferentes respuestas del observado se llega a ellas mediante la abstracción.
- **Unidades moleculares.** Describen atributos observables, con bajo nivel de inferencia y alta carga perceptiva.
- **Unidades temporales.** Estas unidades se refieren al tiempo de intervalo elegido como unidad en los muestreos temporales.
- **Unidades de acción o eventos.** Este tipo de unidades se define por la forma y el contenido. Son similares a las unidades naturales.

Estas unidades no se presentan siempre unívocas, ni la opción por una supone la exclusión del resto, sino frecuentemente relacionadas en el sistema de observación.

3.3.3. Soporte de los sistemas de observación

Los indicadores de la conducta pueden ser expresiones verbales o no verbales. Bar-now (1967) insiste en la necesidad de registrar tanto lo que dice la gente como cómo lo dice, a través de las expresiones del rostro o de las posturas corporales.

Weick (1968) diferencia entre conducta verbal, espacial, extralingüística y lingüística

A) Conductas no verbales

Se consideran conductas no verbales las expresiones faciales, gestuales y posturales (Dollard, 1957).

A.1) Expresiones faciales

Han ocupado un puesto preferente en el estudio de la conducta no verbal y es lógico, puesto que es la más fácil de visualizar y la más significativa por la gran movilidad de la musculatura de la cara. No obstante, su observación y posterior registro requieren un gran entrenamiento de los observadores.

A través de las observaciones faciales podemos distinguir sentimientos, emociones e incluso adquisición o no de conceptos en una situación de aprendizaje.

Leventhal y Sharp (1965) y todos los que han observado los gestos faciales parecen coincidir en que la mitad superior de la cara expresa mejor las emociones negativas, mientras que la mitad inferior proporciona información cuando la emoción es positiva.

La observación de las expresiones faciales exige un gran dominio de los símbolos puesto que a veces los movimientos son muy rápidos y es conveniente registrarlos, aunque su presencia sea de cortísima duración.

Para la codificación de las expresiones faciales, Leventhal y Sharp (1965: 300) elaboraron un sistema de categorías que consiste en una serie de símbolos para registrar la conformidad o disconformidad expresada por:

- Frente y cejas.
- Ojos, párpados y pupilas.
- Boca y nariz.

A cada cinco segundos de observación, para cada apartado, seguían diez segundos de registro. Así, hasta completar los cinco minutos que constituían el período de observación.

A.2) Conductas gestuales

El estudio de las conductas gestuales ha interesado mayormente a la etnografía por su importancia en las diferentes culturas y tradiciones. Brewer (1951) trató el lenguaje gesticular de los árabes levantinos, comparándolo con los americanos. Realizó una lista de gestos simbólicos y gráficos de enorme interés psicológico.

La investigación del gesto realizada con mayor rigor metodológico se debe a Efron (1941), fruto de un estudio de dos años sobre el comportamiento de los judíos de Europa oriental y de los italianos meridionales en la ciudad de Nueva York.

Uno de los gestos más significativos son los realizados con la mirada, e incluso, la evitación de la misma es un indicador de una situación tensa o de ocultación de algo.

A.3) Conductas posturales

La manera como las personas caminan o se sientan puede decirnos mucho acerca de ellas. En la postura influye la tradición cultural, como demostraron los estudios de Hewes (1955) sobre la distribución en el mundo de las costumbres relacionadas con la postura, pero hay indicadores más personales de enorme significatividad para el conocimiento individual. En ellos habrá que distinguir la observación de conductas estáticas (posturas corporales) y dinámicas (movimientos de las diferentes partes del cuerpo).

La observación, de ambas respuestas, implica dificultad de codificación por la variedad de formas que pueden presentar, sobre todo si el individuo observado no tiene restricciones en los movimientos que libremente desee realizar (por ejemplo, en un recreo). En un aula es más fácil realizar estudios estáticos porque el espacio en el que se desplaza el niño es muy pequeño y gran parte del tiempo permanece sentado o en mínimos desplazamientos.

Los movimientos corporales más estudiados han sido los de las manos, y la alteración y equilibrio postural.

Con las manos expresamos multitud de mensajes, reforzamos la opinión expresada a lo largo de un discurso, e incluso transmitimos sentimientos de alegría, rabia, interés, etc. La intensidad de las conductas (acuerdo, enfado, disconformidad, etc.), aunque tengan un soporte verbal, se acentúan o matizan con respuestas gestuales, expresadas por los movimientos faciales y de las manos.

Mead y McGregor (1951) realizaron un estudio con niños balineses y observaron que estos niños aprenden a ajustarse pasivamente a los movimientos corporales de la madre, que los lleva hasta los dieciocho meses en la espalda, de un lado para otro. Se manipulan los miembros del niño para enseñarles a hacer gestos y bailar con las manos antes que aprender a andar. Los dedos inactivos rara vez se hallan en una flexión consecutiva, sino que pueden mantenerse en diferentes ángulos. Cuando un balinés ve luchar a dos gallos, se puede observar, a veces, una crispación de las dos manos, simbólicamente identificadas con cada uno de los gallos (Barnow, 1967).

B) Conducta espacial

La colocación y distancias establecidas por los diferentes grupos humanos no es una disposición fortuita, en ella influye además de la necesidad de cercanía o alejamiento por motivos de trabajo, la afectividad o distanciamiento entre los diferentes interlocutores. Trabajos de gran interés se han realizado a partir de los de Hall (1966), en los que se estudian las conductas espaciales a través de las distancias establecidas entre las personas. «El espacio está estructurado con sorprendente regularidad y cuando se violan los límites espaciales ocurren cambios significativos en la conducta» (Anguera 1985: 104). Esta conducta puede ser observada en sus dos vertientes: estático, referido a las distancias que guardan entre sí los individuos, y dinámico, en lo que habrá que tener en cuenta los desplazamientos que realizan los sujetos en el contexto de observación.

Hall (1978) establece un sistema bastante específico para registrar las conductas espaciales. Distingue cuatro tipos de distancias:

- a) **Distancia íntima.** Entendiendo por ésta el intervalo comprendido entre: una distancia física cercana en la que hay contacto físico, por ejemplo estar con las manos unidas, y una distancia íntima lejana, entendida como medio metro de distancia (aproximadamente) y en la cual aún es posible alcanzar al otro con las extremidades.
- b) **Distancia personal.** Cuando oscila el intervalo de distancia entre medio metro a un metro. Suele ser bastante interesante su observación en la relación profesor-alumno y en los estudios de interacción.
- c) **Distancia social.** Estas distancias mantienen un intervalo de 1 a 3 metros. De 1 a casi 2 metros todavía se entiende como distancia que puede ser dominada por un interlocutor en su relación con el otro. A partir de esta distancia, hasta los tres metros, es una situación de lejanía, observada para registrar colocaciones opuestas.
- d) **Distancia pública,** o intervalo entre el interlocutor más cercano, aproximadamente tres metros, y la distancia máxima a la que llega la voz.

Son numerosos los estudios en el contexto escolar que eligen como objeto de observación la conducta espacial y sus implicaciones en las relaciones sociales. En una investigación sobre ocupación y utilización de los espacios en las zonas de recreo por ambos sexos (Buendía, 1987) se ha confirmado la hipótesis de dominio por parte de los chicos de la mayoría del espacio del campo de juego, el 80 por 100; quedando para las niñas el 20 por 100 restante, que suelen ser colaterales o espacios semicerrados (trinchones, tramos junto a la pared, etc.), a la vez que entre ellas las distancias son más cercanas y constantes que las que mantienen los niños, que pasan de ser de casi íntimas (lucha, forcejeo, etc.) a grandes distancias (corriendo a pillar, escondiéndose, etc.).

C) Conducta extralingüística

Se refiere a todas aquellas conductas cuyos indicadores son vocalizaciones sin contenido semántico. En la conducta vocal se dan diferentes tipos de respuestas, todas ellas de difícil registro por las variaciones que presentan. Con un buen entrenamiento del observador y con medios técnicos se pueden conseguir registros fiables, que permitan conocer estados emocionales o manifestaciones indirectas de los sentimientos de los sujetos.

Cabe distinguir diferentes dimensiones en esta conducta:

- Vocal (intensidad, tono y timbre).
- Temporal (interacción, silencios y sincronización).
- Continuidad (tendencia a interrumpir, dominar, inhibir).
- Estilo verbal (pronunciación, falta de fluidez, lapsus lingüé, etc.).

D) Conducta lingüística

Se refiere a todas las palabras, frases o expresiones con contenido semántico.

Es la más frecuentemente observada, habiendo en la actualidad importantes sistemas de categorías que han permitido codificar con bastante minuciosidad las conductas lingüísticas, tanto en el aula como fuera de ella (Flanders, 1977; Landsheere, 1977, Bales, 1950).

3.4. Criterios para el muestreo observacional

Hay dos criterios que definir antes de realizar el muestreo:

- a) Selección del segmento de conducta (evento *versus* estados).
- b) Sujeto o sujetos sobre los que recae la observación.

3.4.1. Conducta momentánea versus estados

A través de las técnicas de muestreo seleccionamos los segmentos de la conducta que nos interesan estudiar, siempre que éstos sean representativos del total. Cabe preguntarse: ¿por qué hay que fraccionar una manifestación conductual y no estudiar todo el continuo, y evitaríamos la duda de su representatividad? Esto es evidente, y de hecho, cuando se puede, debe realizarse, pero no siempre es así. Por ejemplo: en una investigación sobre las interacciones entre iguales en la realización de una tarea, el estudio se debe hacer sobre toda la secuencia interactiva; pero si lo que interesa es conocer cómo una madre interactúa con su hijo de cuatro años, la cantidad de datos que se recogerían de un solo día, minuto a minuto, sería enorme, por lo que tendríamos que comenzar acotando situaciones o tiempo. ¿Podríamos decir que con media hora diaria durante diez días se recogen todas las posibles situaciones? ¿Sería preferible observar sólo las conductas significativas cuando éstas se producen? ¿Sería más representativa la conducta interactiva de diez parejas, madre-hijo, observadas una vez cada día?

Para estas y otras preguntas necesitamos recurrir a diferentes técnicas de muestreo según que el objeto de observación sea: una situación de evento (conducta fugaz) o una situación de estado, que implica una cierta continuidad (Anguera, 1988).

Si en un Centro de Educación Primaria se desea estudiar la solidaridad en un grupo a través de las ayudas que se facilitan los alumnos, podemos plantear, entre otros, los siguientes interrogantes:

- ¿Cuántas veces ante una petición de ayuda se responde con ayuda?
- ¿Son más solidarios y participativos los alumnos que las alumnas durante el desarrollo de una tarea?

En situaciones de aprendizaje en grupo, ¿cuántas veces ayuda a los compañeros? En la primera pregunta nos interesa sólo una conducta momentánea. El interés se centra en la aparición o no de la conducta de ayuda.

La segunda es una situación de estado. Habrá que registrar no sólo el número de ayudas, sino también el período de duración de las mismas; son situaciones en las que interesa tanto el inicio de la conducta como el final de la misma, e incluso el desarrollo de la secuencia.

En la tercera se vuelve a estar interesado sólo en una conducta de evento; la frecuencia con que se inician ayudas. Es igualmente una observación momentánea y como tal sólo interesa registrar la aparición de la misma.

3.4.2. Elementos de la muestra

Otro de los inmediatos interrogantes que surgen al hablar de muestreo es el número de individuos que tienen que formar la muestra.

Cuando la observación se realiza en una investigación experimental, como técnica de recogida de datos, la selección de la muestra y el número de sujetos que la forman siguen el criterio del procedimiento experimental. El objetivo, en la selección de la muestra, es que sea representativa de la población, para poder extrapolar los resultados obtenidos a la población de referencia. La observación, en estos casos, se puede realizar sobre grupos formados aleatoriamente, sobre un grupo elegido al azar o sobre un sujeto, diseños de caso único (Suen y Ary, 1989). Pero si hablamos del método de observación sistemática, las dimensiones que subyacen en el muestreo son diferentes. No se buscan muestras representativas puesto que el objetivo no es generalizar los resultados a ninguna población.

La metodología observacional es por naturaleza ideográfica (Anguera, 1988), por lo que el foco de atención será un individuo o, como mucho, un pequeño grupo que funcionan como unidad para el objetivo de la observación. Por ejemplo, en el estudio de la interacción entre iguales, el foco de atención pueden ser dos, o incluso tres individuos; pero aun así, no tiene el carácter nomotético de la metodología experimental y por tanto no hay por qué buscar la representatividad del grupo.

3.5. Clases de muestreo

En la literatura sobre el tema no existe acuerdo sobre las diferentes técnicas de muestreo en la metodología observacional, y una de las razones es, como apunta Anguera (1990), el diferente criterio seguido a la hora de realizar la clasificación. Fassinacht (1982: 119) cuando habla de técnicas de muestreo, realmente se está refiriendo a las técnicas de registro, que implican decisiones de cómo debe registrarse la observación. Distingue básicamente dos métodos de muestreo: Muestreo de tiempo y muestreo de eventos, y en cada uno especifica unas técnicas para su cuantificación.

Por su parte Anguera, en las sucesivas publicaciones, siempre toma como referencia el criterio de «cuándo» se debe observar, y «qué o a quién» se debe observar.

Nosotros los clasificaremos en función del mayor o menor grado de sistematización de la situación de muestreo. Según éste, se puede realizar un continuo que va

desde el menos sistematizado muestreo *ad libitum* (Altmann, 1974) y bajo control externo, hasta el más sistematizado que correspondería al muestreo temporal, con un alto nivel de control externo. En el espacio intermedio del continuo situamos el muestreo focal y el muestreo de todas las ocurrencias de algunas conductas.

3.5.1. Muestreo ad libitum

Es un muestreo utilizado fundamentalmente en la observación exploratoria; de ahí su carácter informal y asistemático. Su principal misión es proporcionar información sobre una situación y facilitar la sistematización posterior del estudio. En este tipo de muestreo se recoge todo lo que llama la atención sin que el objetivo esté previamente establecido ni delimitado.

Son las notas de campo, propias de la observación participante, y que presenta el inconveniente del sesgo del observador, tanto en la selección de las conductas como en la atribución de la idéntica probabilidad de aparición de todas ellas. En un muestreo *ad libitum* se registra igualmente una conducta rara y puntual que una manifestación es normal en el desarrollo de las diferentes sesiones. Por ejemplo, en una sesión de observación puede haber razones para que un niño sociable se comporte aislado o no participativo, esta conducta se recoge por igual que la del niño que normalmente es solitario o inadaptado.

Mediante este muestreo rara vez es posible determinar qué diferencias en los datos se deben a diferencias verdaderas en los individuos, y cuáles a meros sesgos en el muestreo (Altmann, 1974).

Son registros de gran utilidad en las primeras fases exploratorias, como material ilustrativo de la conducta que interesa estudiar.

Quera (1986) realizó un microanálisis de la conducta interactiva aplicado a la interacción materno-filial en chimpancés; en este estudio, los tres primeros días de observación, de un total de 16, realiza el muestreo *ad libitum* con dos finalidades. Primero realizar una categorización de las conductas interactivas, y segundo, como sesiones de entrenamiento para los observadores. Posteriormente, tras dos sesiones de observación semisistemizadas, en las que pone a prueba las categorías establecidas, procede a la obtención de datos con un muestreo focal (que veremos a continuación).

En estas situaciones es indudable el valor de este tipo de muestreo, como paso previo al muestreo sistemático, único procedimiento para evitar el sesgo que resulta cuando la atención del observador es atraída por ciertas clases de conductas o de individuos (Altmann, 1974).

3.5.2. Muestreo focal

Se denomina muestreo focal cuando en períodos de tiempo, previamente establecidos, registramos todas las acciones de un individuo o de un grupo de individuos que previamente han sido seleccionados para ser el «foco» de observación.

Una vez seleccionado, el observador registra la secuencia de conducta del grupo, mientras que dure el período de tiempo establecido. Una variante es la observación de cada miembro del grupo, durante un tiempo determinado, por la división del intervalo total de observación en partes proporcionales a los sujetos que componen el grupo.

Aunque aleatoriamente se elige al sujeto por donde iniciar la observación, posteriormente se puede ir rotando entre los miembros del grupo.

Para evitar la posible no equivalencia de las muestras, por incidentes distorsionantes o desigual aparición de la conducta, se puede proceder a nuevas particiones del intervalo de tiempo, de tal manera que cada sujeto sea observado en más ocasiones, aunque en períodos de tiempo más cortos (Anguera, 1988).

No obstante, este tipo de muestreo es preferible utilizarlo para un único individuo focal, a lo sumo sobre una pareja, sobre todo si se trabaja en condiciones de observación no totalmente perfectas. Por ejemplo, el intento de realizar un muestreo focal sobre un grupo de tres alumnos-amigos en un recreo sería difícil por la incompatibilidad de captar situaciones simultáneamente, a veces distantes entre sí y fuera del alcance del observador.

3.5.3. *Muestreo de todas las ocurrencias de algunas conductas*

Este muestreo tiene un indudable valor, cuando lo que se pretende observar es la aparición de una determinada conducta.

Las condiciones que se le exigen son tan rigurosas que no siempre es posible su utilización. Como señala Altmann (1974), este tipo de muestreo es útil siempre que:

- Las condiciones de observación sean excelentes.
- Las conductas atraigan la atención suficientemente como para que puedan ser observadas todas las cosas.
- Los eventos conductuales no ocurran en frecuencias muy altas.

Si se cumplen estas condiciones, se puede recoger importante información sobre la tasa de ocurrencia de conductas, e incluso registrar conductas simultáneas cuando el objeto de estudio es la sincronía conductual.

Cuando todos los participantes del grupo pueden ser identificados en cada ocurrencia de la conducta, esta técnica sería equivalente al muestreo focal, referido al grupo-focal.

La información obtenida con este muestreo, además del interés que aporta por sí misma, puede ser una herramienta importante para probar supuestos que se establecen al utilizar otras técnicas de muestreo.

3.5.4. *Muestreo temporal*

Es un muestreo de la conducta de un individuo durante unos períodos de tiempo cortos y previamente definidos y el registro, dentro de dichos períodos, de la ocurrencia o no ocurrencia de ciertas formas de conducta especificadas y definidas objetivamente.

El término muestreo temporal ha sido utilizado de diversas maneras por diferentes autores, e incluso se ha aplicado a cualquier muestreo en el que se mantiene una unidad fija de observación (Fassnacht, 1982).

Las características generales de esta técnica, siguiendo a Altmann (1974), son las siguientes:

- a) En cada período de observación se registra la ocurrencia o la no ocurrencia de la conducta.
- b) Al tratarse de un muestreo de secuencias, no de eventos, la ocurrencia significa proceso de cierta duración en el período de muestreo.
- c) Los períodos suelen ser cortos (no superiores a 15 o 20 segundos) y suelen repetirse, muchos, sucesivamente.

En este apartado, cuando hablamos de muestreo temporal, nos referimos a presencia o ausencia del rasgo observado en un punto del tiempo (muestreo instantáneo), y a muestreos en intervalos temporales.

A) *Muestreo instantáneo*

El muestreo instantáneo, o de puntos de tiempo, implica observar, y consecuentemente registrar, cada cierto tiempo, que previamente ha sido determinado, la presencia o no de la conducta en ese momento. Mediante una señal acústica se marca el punto en el tiempo que debe realizarse la observación. Es de utilidad cuando se pretende registrar el porcentaje de tiempo que los individuos dedican a una actividad, puesto que dicho porcentaje se estima a partir del porcentaje de muestras en las que se ha realizado el registro (Smith y Connolly, 1972). Para esto es necesario que la distancia entre los puntos de tiempo sea pequeña, sin llegar a extremar este aspecto que haría perder las virtudes del muestreo, transformándolo en registro continuo.

B) *Muestreo por intervalos*

En esta modalidad la observación se realiza durante un espacio temporal (por ejemplo cinco minutos), transcurridos los cuales se realiza el registro, anotando sólo la presencia o ausencia de la conducta en ese intervalo. Este muestreo presenta el problema de la delimitación de la longitud de los intervalos; decisión que hay que determinar por tanto, en función de cada caso de estudio.

Aunque este método se ha utilizado bastante en las últimas décadas, tanto en estudios sobre conducta animal (Rhine y Flanigon, 1978) como en conducta humana (Wright, 1960; Hutt, 1974; Richards y Bernal, 1972), no está exento de críticas; sobre todo en lo referente a la utilización de estas puntuaciones como medida del tiempo invertido en una secuencia de conducta, o la frecuencia de los actos interactivos. Altmann (1974) puntualiza en este sentido, que una puntuación uno-cero no es la frecuencia de conducta, sino la frecuencia de intervalos que incluyen cualquier cantidad de tiempo invertido en la conducta; y segundo, que no es cierto que el porcentaje de intervalos sea lo mismo que el porcentaje de tiempo invertido en una actividad.

Por ejemplo, si el llanto de un niño se inicia en un intervalo, continúa durante todo el segundo intervalo y finaliza en el tercero, esto se puntúa con un tres; una unidad por cada intervalo. Pero si rompe a llorar tres veces en un intervalo y nada en los otros dos, la puntuación resultante será 1.

Es indudable que el muestreo temporal también tiene importantes ventajas, entre las que cabe señalar la facilidad de las observaciones. Por ejemplo, si dos observadores están observando la conducta de llanto de un niño, es evidente que al no tener que determinar las veces que ocurre, sino sólo la presencia o ausencia de esa conducta durante cada intervalo, el índice de coincidencia es casi total. Pero coincidimos con Altmann que un mayor acuerdo no garantiza más información; por lo que creemos que siempre que sea posible el muestreo temporal se debe utilizar conjuntamente con otros métodos, descritos anteriormente, con lo que ganamos en precisión e información.

4. DISEÑOS EN LA INVESTIGACIÓN OBSERVACIONAL

El diseño es un plan estructurado de acción que, en función de unos objetivos básicos, está orientado a la obtención de datos relevantes para los problemas y cuestionamientos planteados.

La estructura básica de los diseños observacionales, en lo que se refiere a su tipología, se definen básicamente por el cruce de dos dimensiones bipolares: nomotético-idiógráfico y sincrónico-diacrónico (Anguera, 1990), distinguiéndose esencialmente dos tipos de diseños: secuenciales y transversales.

4.1. Diseños secuenciales

El objetivo de un diseño secuencial es el estudio de manifestaciones momentáneas o sucesos de conducta de los sujetos, bien para analizar la estabilidad de dicha conducta, o bien las relaciones existentes en la conducta de sujetos en interacción.

Esto permite establecer dos tipos de diseños secuenciales, cuya característica común es el seguimiento en el tiempo:

- a) Diseños de caso único y seguimiento.
- b) Diseños de grupo y seguimiento.

Los diseños secuenciales implican una forma de proceder a lo largo de todo el proceso investigador, determinando ciertas condiciones en todas las fases de dicho proceso. Así, desde la formulación del problema, todos los pasos están orientados a obtener una información adecuada para el estudio secuencial de la conducta. Los datos deben proporcionar, al menos, información sobre frecuencia y orden de aparición de las conductas categorizables.

Los diseños de caso único y seguimiento son los más indicados en la metodología observacional por su carácter esencialmente ideográfico, y se registra frecuencia de cada conducta, orden de aparición, y a veces tiempo.

Por ejemplo, para conocer las mediaciones del profesor en la resolución de tareas aritméticas, se realiza un seguimiento de las intervenciones en los diferentes contenidos durante un tiempo determinado. Se podrían representar como sigue:

- Categorías observables:

A: facilita información
B: plantea preguntas
C: responde a las dudas
D: anima a intervenir

- Sujeto observado: X.

- Registro realizado:

A D A B C A B D C B D A B C A B C D A D B C A C D C D

Bakeman y Quera (1996), en la presentación del programa SDIS, para el análisis de datos secuenciales, plantean cuatro categorías o tipos básicos de datos:

1. Datos secuenciales de evento.
2. Datos secuenciales de estado.
3. Datos secuenciales de evento con tiempo.
4. Datos secuenciales de intervalo.

El primer tipo sólo contiene información sobre los eventos.

Por ejemplo: A B C. Indica que el evento A fue seguido por el B, que fue seguido por el C.

Se utilizó este diseño en un estudio de interacción de una pareja. Los observadores extrajeron las siguientes unidades de conducta para codificar las interacciones verbales (son códigos mutuamente excluyentes y exhaustivos):

1. Se queja.
2. Actúa de forma emotiva.
3. Aprueba.
4. Muestra empatía.
5. Niega.
6. Otra.

El segundo tipo, al ser situaciones de estado, implica recoger información sobre el tiempo o duración. Por ejemplo: A,10:30 B,10:42. Indica que la conducta A se inició a las 10:30 y deja de ocurrir a las 10:42 que se inicia la conducta B. Estos diseños incluyen los datos secuenciales de estados múltiples, por ejemplo A,10:30 B,10:42 y C,10:30 D,10:36; lo que indica que hubo un segundo flujo de conducta coocurrente con el primero y que ambos se iniciaron a las 10:30.

Este diseño se utilizó (Adamson y Bakeman, 1984) para describir el desarrollo de la atención conjunta en niños mientras interactúan con distintos compañeros. Se graba a cada niño diez minutos durante diferentes períodos de tiempo. Los observa-

dores segmentan la conducta grabada en una serie de estados de atención. Se categorizó de la siguiente manera:

- 1. Desocupado: el niño no atiende.
- 2. Observador: el niño observa las acciones de otro, pero no toma parte en ellas.
- 3. Persona: el niño juega cara a cara con una persona, pero sin emplear objetos.
- 4. Objeto: el niño juega con objetos, pero no con una persona.
- 5. Conjunta: el niño comparte con una persona la atención hacia un objeto.

Los códigos, al igual que en el estudio anterior, son mutuamente excluyentes y exhaustivos.

El tercer tipo incluye tanto conductas momentáneas como duraderas, lo que puede provocar solapamientos. Por ejemplo: A,10:30-10:42 B,10:33 C,10:37-10:54 B,10:50; lo que indica que la conducta A se inicia a las 10:30 y dura hasta las 10:42 mientras que el evento B fue momentáneo, etc.

Este diseño fue utilizado en un estudio para evaluar el efecto de la conducta de los adultos en el malestar infantil durante intervenciones médicas dolorosas (Manne, Bakeman, Jacobson, Gorfinkle, Bernstein y Reed, 1992). Se registraron los tiempos de comienzo para todas las conductas y además los de finalización para las duraderas. Las conductas fueron:

- 1. El niño resiste (duradera).
- 2. El niño manifiesta un malestar momentáneo.
- 3. El niño llora (duradera).
- 4. El adulto explica el procedimiento al niño (duradera).
- 5. El adulto distrae al niño (duradera).
- 6. El adulto cede el control al niño.
- 7. El adulto alaba al niño.

Estos códigos no son mutuamente excluyentes y en realidad interesaba conocer qué códigos coocurren y durante cuanto tiempo.

El cuarto tipo implica acotar intervalos y representar las ocurrencias de las conductas en dichos intervalos. Por ejemplo: AB,A,BCD; lo que indica que la conducta AB ocurrió en el primer intervalo, la A en el segundo y la BCD en el tercero.

El diseño fue empleado por Konner (1996, 1997) para conocer el desarrollo de los niños IKung y la interacción con sus semejantes. Se observaron niños de diferentes edades empleando un sistema de códigos previamente elaborado. Cada niño fue observado seis sesiones durante 15 minutos en diferentes momentos del día. Algunos de los códigos fueron:

- 1. El niño ofrece objetos.
- 2. El niño juega con un objeto.
- 3. El niño ríe o sonríe.
- 4. El niño vocaliza.
- 5. Otra persona ofrece objetos al niño.
- 6. Otra persona alienta al niño.
- 7. Otra persona entretiene al niño.

Estos códigos no son mutuamente excluyentes ni exhaustivos.

4.2. Diseños transversales

En los diseños transversales se plantea la relación entre diversas variables de estudio. Por ejemplo, si se analizan diferentes dimensiones en la interacción de dos sujetos, a través de la aplicación de dos sistemas de categorías, se obtiene una clasificación cruzada del proceso interactivo.

Dimensiones de la interacción

Sistema de categorías	X	Y
A (a ₁ , a ₂ , a ₃ , ...)	(a ₁ , a ₂ , a ₃ , ...)	X
B (b ₁ , b ₂ , b ₃ , ...)	(b ₁ , b ₂ , b ₃ , ...)	X
		(b ₁ , b ₂ , b ₃ , ...)
		Y

Los dos sistemas (A y B) han sido aplicados concurrentemente a los dos elementos o dimensiones fundamentales (X e Y).

5. EL CONTROL DE LOS SESGOS EN LA OBSERVACIÓN

En la metodología observacional, al realizarse en contextos naturales, la exigencia de control de las situaciones que pueden alterar los resultados es uno de los requisitos fundamentales en todo el proceso.

Los sesgos pueden ser ocasionados por factores situacionales, personales y metodológicos.

5.1. Control de los sesgos contextuales

La definición del contexto, en el que tiene lugar la observación, permite al investigador interpretar las observaciones y darles el significado preciso que la conducta tiene en el momento de producirse.

McGaw, Wardrop y Bunda (1972) consideran que estos factores situacionales han sido olvidados en los diseños observacionales, argumentando que la varianza en la conducta del profesor y del alumno, de una situación a otra, puede ser más importante que la simple consideración del azar en la variabilidad de estas situaciones.

Entendemos por factores situacionales: la materia o disciplina sobre la que se está trabajando, el tamaño de la clase, la estructura del grupo, el momento del día o de la semana, etc.

El control de estas situaciones se puede realizar con una buena descripción del contexto como requisito previo en la observación, o diferenciarlas en el diseño como fuentes de variación, diferentes, por ejemplo, de la observación del profesor o del alumno. Frick y Semmel (1978) consideran que la opinión de McGaw, Wardrop y Bunda (1972) debe ser tenida en cuenta cuando se diseña un estudio observacional, y se evitaría atribuir al error de medida la varianza que debe ser atribuida a las situaciones.

5.2. Control de los sesgos personales

Dentro de estos factores vamos a considerar al observador y al sujeto observado.

5.2.1. Sesgos ocasionados por el observador

El observador, cuando percibe la conducta o cuando la dota de significado, puede cometer sesgos que deben ser controlados para un fiel registro e interpretación de las observaciones.

En cuanto a la percepción de la conducta, ésta se puede alterar tanto por una inadecuada atención como por las deformaciones ocasionadas por la memoria, si el registro no se hace momentáneamente. Igualmente, el observador puede verse influido por el conocimiento que tiene de la conducta observada, o del sujeto objeto de la observación, apareciendo el sesgo de expectancia, o precipitación del observador en registrar rasgos de conducta que no aparecieron durante la observación.

Estos errores se pueden controlar utilizando cámaras de vídeo y grabadoras que ayuden a una perfecta reproducción de la conducta observada.

Cuando se dota de significado la conducta percibida, pueden cometerse otra serie de errores, siendo los más frecuentes:

- Olvido o desconocimiento de la intencionalidad de la conducta humana.

Con frecuencia el observador atribuye el mismo significado a rasgos que aparentemente son idénticos, olvidando la propositividad de la conducta humana y dando el mismo significado a conductas generadas en contextos diferentes.

El conocimiento es situado, y un mismo acto puede tener significados diferentes dependiendo del contexto en el que se produce.

- Proyección de la personalidad del observador, identificándose con situaciones e ignorando la aparición de otras que a veces involuntariamente pasan desapercibidas. Esto puede estar influido por su estado de ánimo o por sus actitudes y valoraciones.

Una buena formación de los observadores y adecuados medios técnicos, como comentamos anteriormente, evitarían en gran medida estos sesgos.

5.2.2. Sesgos ocasionados por los sujetos observados

El más importante de los sesgos es la propia presencia del observador. Al realizar la observación en el aula los alumnos perciben la presencia de un desconocido, alterando la normalidad de la conducta hasta que se acostumbren a la permanencia en la clase de un extraño. Una forma de controlarlo es prescindiendo de las primeras sesiones de observación hasta que vuelva a normalizarse la conducta.

5.3. Control de los sesgos metodológicos

En el diseño de la metodología observacional pueden cometerse errores que afectan fundamentalmente a la fiabilidad de las observaciones y al contraste de las hipótesis de investigación. Algunos aspectos relevantes en los diseños observacionales son:

- *Asignación de observadores.* Un tema debatido en metodología observacional es el número de observadores por clase que deben registrar las conductas. Medley y Mizel (1958) encuentran que el número de visitas por clase aumentaba más la fiabilidad de las observaciones que el incremento del número de observadores en cada sesión, aunque posteriormente, en 1963, consideran que la situación ideal sería que pares independientes de observadores realicen observaciones en cada clase, en un número igual de sesiones. En cualquier caso, un buen diseño de los momentos y sesiones de observación, y que sean realizadas por más de un observador, aseguran mayor concordancia en los registros y un aumento de la fiabilidad de las observaciones.

- *Pobres diseños de los sistemas de observación.* La elaboración de sistemas de categorías, con incompletas definiciones o basadas en indicadores extremadamente complejos, son causa de graves problemas en los registros y aún más en la concordancia de los observadores. Medley y Norton (1971) consideran que para evitar estos errores las categorías deben ser fácilmente reconocibles y lo menos dependientes posibles de conocimientos sofisticados o de los propios valores de los observadores.

Herbert y Attridge (1975) presentaron unos criterios para el desarrollo de diseños observacionales que minimizan los errores metodológicos. Dichos criterios se basan en la identificación, validación y practicabilidad de los sistemas de observación. No obstante, la propia elaboración de los sistemas de categorías, acompañada de una buena preparación de los observadores, minimiza los errores metodológicos al evitarse la extrapolación de sistemas generados en contextos diferentes.

6. REGISTRO DE LAS OBSERVACIONES

En el momento en que el observador comienza a obtener datos para la falsación de la hipótesis planteada es necesario un registro para que la realidad que percibe quede plasmada de la forma más clara posible para su análisis posterior.

6.1. Clases de registros observacionales

Hay diferentes procedimientos e instrumentos para registrar las observaciones, como exponemos a continuación, pero el registro por excelencia de la observación sistemática es el sistema de categorías.

6.1.1. Registro anecdótico

El registro anecdótico o anecdotario permite recoger observaciones que de no registrarlas con este procedimiento pasarían desapercibidas o se deformarían en el recuerdo por descontextualización u olvido.

Habitualmente los registros se hacen en una ficha de cartulina en la que figura: el hecho tal y como ocurrió, fecha, hora y cuantos datos puedan identificar más exactamente la situación que originó ese comportamiento. Detrás se acompaña la interpretación personal de quien la observó. Esta interpretación no debe dilatarse en el tiempo porque podrían falsearse las observaciones, ni tampoco realizarla inmediatamente que sucede, porque los aspectos que rodean la conducta que se registra han podido influir en el observador e incitarle a realizar una interpretación errónea de la misma.

Cada ficha debe registrar un incidente o anécdota. El conjunto de fichas realizadas a la largo de un período de tiempo formará el anecdotario sobre una persona.

Este fichero, de larga elaboración, es de singular importancia para el estudio del comportamiento social. Es en el aula donde más tiempo pasa el niño entre sus iguales y es, por tanto, donde mayor riqueza de información podemos obtener.

Los registros deben estar actualizados. La conducta del alumno cambia y no son válidas las extrapolaciones en el tiempo de los registros observacionales. Así pues, debe realizarse sobre aquellos sujetos que van a ser objeto de seguimiento en la investigación y se anotarán sólo aquellos incidentes relacionados con las conductas que nos interesan estudiar y que realmente sean significativas.

Modelo de ficha anecdótica

Alumno _____	Fecha _____
Curso _____	
Actividad que se estaba realizando _____	
Lugar _____	
Anécdota:	
Jugando con un grupo de cinco amigos al corro, puso la zancadilla a una niña del grupo, pasando posteriormente a justificar su acción, argumentando que había sido sin querer.	
Interpretación:	
Luisa es una niña acostumbrada a ser de las primeras en la clase. A la niña que le ha puesto la zancadilla es sobresaliente en coordinación motora y muy sociable. El acto ha parecido involuntario, pero suelen pelearse con frecuencia.	

A veces no es necesario realizar un fichero con una ficha para cada anécdota y es suficiente un cuaderno en el que se anoten de manera ordenada los incidentes de interés. Ejemplo:

Fecha	Lugar	Incidente
24-II-1988	Recreo	_____
»	»	_____
»	»	_____

Las interpretaciones pueden ir en un cuaderno paralelo en el que se anota:

Fecha	Lugar	Interpretación del incidente
24-II-1988	Recreo	_____
»	»	_____
»	»	_____

A) Limitaciones de los anecdotarios

Siguiendo el trabajo de Buendía (1992) enumeramos a modo de síntesis los principales inconvenientes de este instrumento, cuando su uso o interpretación no se hace correctamente:

1. Anotar no el incidente, sino su interpretación.
2. Usar el anecdotario en defensa propia del que lo ha escrito.
3. Anotar incidentes sin dejar constancia del contexto o situación en que tuvo lugar, cuando este contexto es interesante.
4. Juzgar a un individuo por unos pocos incidentes que no representan lo que realmente es.
5. No guardar la información como material confidencial puede perjudicar al sujeto en el futuro.
6. Dejarse impresionar por los acontecimientos negativos más que por los positivos y apenas dejar constancia de estos últimos.
7. Dejarse llevar por juicios preconcebidos, y seleccionar, para bien y para mal, lo que debe ser anotado.

6.1.2. Lista de rasgos

Consiste en un listado de operaciones, o secuencias de acción, que el investigador utiliza para registrar su presencia o ausencia como resultado de una atenta observación. Las características que deben reunir son:

- Los aspectos que van a ser observados deben plantearse de manera clara y concisa.
- La mayoría de las listas sólo admitten presencia o ausencia del rasgo a observar.
- La presencia o ausencia del rasgo no debe interpretarse como una forma de medición, sino sólo como una información descriptiva de la conducta observada.
- Cuando la lista de rasgos se realiza para observar un proceso secuencial, las conductas reflejadas deben presentarse ordenadas y en el mismo sentido en el que aparecen las secuencias a las que representan.

6.1.3. Escalas de estimación

Con estas escalas, también denominadas de puntuación o calificación, conseguimos registrar no sólo la presencia o ausencia del rasgo, sino también el grado o intensidad con el que el observador percibe la presencia de dicho rasgo. Como consecuencia estas escalas, aun siendo de gran utilidad para el docente, tienen el inconveniente de la subjetividad que conlleva la emisión de un juicio; por lo que es aconsejable que intervengan varias personas en la calificación o estimación de los rasgos a evaluar.

En la elaboración de una escala hay que tener en cuenta los siguientes requisitos:

- Especificar con claridad los objetivos que se intentan verificar.
- Seleccionar los rasgos más sobresalientes de la conducta a evaluar y que a la vez sean independientes entre sí.
- Establecer unidades de observación sobre conductas claramente observables y verificables.
- La unidad de observación estará en función de la cantidad de rasgos aislados como necesarios para dar respuesta al objetivo de la escala.
- En la evaluación debe intervenir más de una persona para evitar el sesgo del observador y conseguir una valoración más precisa.

A) Tipos de escalas

Hay tres clases de escalas de puntuación.

a) Escala numérica

La intensidad o grado de la conducta observada se representa por un número, que oscila entre 0 y 5. Previamente se establece la correspondiente equivalencia entre el valor numérico y el grado de presencia de dicha conducta.

Ejemplo: Con la siguiente escala se pretende valorar algunos aspectos de la asignatura X.

Las equivalencias entre criterios y calificaciones son:

5. Muy suficiente,	Muy grande	Excelente
4. Suficiente,	Grande	Bueno
3. Regular,	Mediano	Regular
2. Insuficiente,	Pequeno	Baja
1. Muy insuficiente,	Muy pequeno	Muy Baja

Indica de 1 a 5 el grado en el que se han conseguido los siguientes aspectos:

- Satisfacción global obtenida en la materia.
- Relación personal facilitada por el profesor.
- Interés en enseñar del profesor.
- Evaluación de tu interés a lo largo del curso.

b) Escalas gráficas

En las escalas gráficas, el profesor señala mediante una cruz, o línea horizontal, la categoría que mejor responda al rasgo de conducta o característica evaluada.

Ejemplo: Con la siguiente escala se pretende medir si cumplen los alumnos ciertas normas de clase.

	Siempre	Casi siempre	A veces	Nunca
Llega puntualmente a clase	×		×	
Respeto el material de los compañeros				×
Entrega los trabajos en el tiempo convenido	×			

c) Escalas descriptivas

En estas escalas, la descripción del rasgo o característica es mayor que en las anteriores, con lo que se evita que el observador asigne un significado personal a la conducta observada.

Ejemplo: Valoración del sentimiento de cooperación. Capacidad para trabajar en equipo con los compañeros y superiores (Lafourcade, 1972: 168).

	10-9	8-7	6-5	4-3	2-1
Siempre dispuesto a prestar su ayuda		Evidencia conformidad al integrar grupos de objetivos vinculados a su labor docente. Superpone el éxito del trabajo colectivo al suyo personal.	Coopera, pero sin mayor esfuerzo y sin mostrar mucha voluntad en ello.	Prefiere trabajar solo.	No presta ni permite ayuda. Evita todo trabajo en común.

Cada categoría describe un rasgo de conducta. Señala la que a tu juicio mejor responde a la realidad del alumno observado.

6.1.4. La categorización

El instrumento por excelencia de recogida de informaciones es el sistema de categorías. Una categoría es una conceptualización realizada a partir de la conjunción de elementos concretos que tienen características comunes.

La creación de un sistema de categorías supone conceptualizar previamente cada una de ellas, de tal manera que la inclusión de un elemento en una categoría elimine la posibilidad de inclusión en cualquier otra, y a la vez cualquier elemento definido como objeto de observación debe estar representado en la conceptualización de alguna de ellas.

Esto exige unas reglas de categorización que exponemos a continuación:

1. Las categorías deben definirse de manera clara y precisa para que la adscripción de los rasgos se realice con el menor error.
Los criterios seguidos para la conceptualización pueden ser muy diversos: nivel de molecularidad-molaridad (Sackett, 1978), su nivel de apertura (Anguera, 1990) o la borrosidad de sus límites (Mervis y Rosch, 1981). En función de la elección de unos u otros se generan categorías con diferentes características de definición.
2. Deben ser exhaustivas. Cualquier comportamiento del ámbito considerado como objeto de estudio pueda asignarse a una de las categorías.
3. Mutuamente excluyentes. La asignación de un caso a una categoría impide la asignación a otra. El principio de la mutua exclusividad no es a veces posible, siendo necesario la creación de categorías de conductas coocurrentes, e incluso categorías de límites poco claros o categorías borrosas (McCloskey y Glucksberg, 1978).

6.1.5. Selección y definición de las unidades de observación

Elegir y adaptar o, lo que es más frecuente, crear un sistema de categorías, se basa en una sucesión de actividades complementarias, a través de las cuales el observador decide qué información es relevante de acuerdo con los objetivos de la investigación.

En el proceso de categorización se seleccionan las conductas pertinentes, se toman decisiones en torno a los criterios de inicio y final de las mismas, y se definen los elementos que formarán parte de una determinada categoría (comparten un mismo código), especificando cuáles son los atributos comunes y cuáles los diferenciadores de los elementos pertenecientes a las otras categorías. Así pues, una de las más importantes decisiones es la selección de la unidad de conducta.

Por ejemplo, Als, Tronich y Brazelton (1979) describieron la interacción existente entre un niño y su madre, teniendo como finalidad el hacer que el niño juegue con ella. Los autores dividen esta tarea en cuatro pasos de secuencia que son: la iniciación, la orientación mutua, el saludo y el diálogo. Cada uno de estos pasos es descrito luego con los diferentes términos que lo componen. Por ejemplo, para la madre, la iniciación consiste en mostrar cara animada, lo cual incluye abrir los ojos, mover las cejas, enfocar los ojos hacia el niño, vocalizar fuerte, etc.

Si la finalidad es describir la frecuencia con que las madres motivan a los hijos para jugar con ellas, serían suficientes las cuatro categorías, pero si el interés del investigador es enseñar a una cuidadora la tarea que realiza la madre al jugar con el niño, el nivel de desglose de los componentes de cada tarea es mucho mayor; sería conveniente describir todos los indicadores de relación, expresión, vocalización... para que la similitud fuera máxima (McFall, 1982). El nivel apropiado dependerá de los intereses del investigador, y guiar un ojo o levantar un dedo puede ser importante en una investigación y pasar totalmente desapercibido en otra si la unidad de conducta elegida es más general.

En función del problema se tendrá que tomar la decisión en la elección de la unidad de conducta. Hawkins (1982) señala la conveniencia de codificar comportamientos moleculares, puesto que más tarde se pueden combinar en medidas molares; mientras que a la inversa es imposible. Altmann (1965) señala que las unidades de conducta presentan los problemas básicos de cuándo dividirse y cuándo agruparse.

En esto subyace el problema del *continuum* establecido entre molaridad y molecularidad (Sackett, Ruppenthal y Gluck, 1978). La molaridad presenta la ventaja de observar conductas con mayor significado, pero pierden el valor objetivo de las conductas moleculares, sobre todo cuando el estudio exige grados relativamente bajos de interpretación e inferencia. No obstante, siempre se puede comenzar con unidades concretas que posteriormente pueden agruparse, para dar lugar a niveles de análisis más generales.

Una vez decidido el nivel de conducta elegido, la descripción de dicha unidad se presenta como requisito imprescindible para su operativización y análisis posterior. La descripción de la unidad de observación se puede realizar operativa o conceptualmente. Una definición operativa supone asignar significados a la unidad de observación, especificando las actividades que hay que realizar para su catalogación. La definición conceptual supone explicar la esencia de la unidad de conducta observada.

6.2. Procesos de elaboración de un sistema de categorías

Un sistema de categorías debe ser una excelente representación de la realidad que interesa observar y que dependerá de la buena definición que de las categorías se haga, su mayor o menor correspondencia con esa realidad.

Hay tres procesos diferentes en la elaboración del sistema de categorías: deductivo, inductivo y deductivo-inductivo.

En el primero se parte de un marco teórico para la conceptualización y amplitud de las categorías.

En el segundo se parte de registros narrativos, cuadernos de campo, o registros realizados con vídeo o magnetofón en periodos de observación exploratoria; y a partir de ese material, se extraen los rasgos que serán agrupados en función de la similitud de ciertas características, pertinentes al objeto de investigación. La conceptualización última de los rasgos afines, tras sucesivos pasos de filtrado, análisis, elaboración y síntesis, configurará el sistema de categorías.

El tercer proceso seguido para la elaboración de sistemas de categorías es deductivo-inductivo. Se parte de un marco teórico para definir las macrocategorías y posteriormente se procede a la elaboración de listas de rasgos extraídos a partir de los registros que se realizan en el contexto natural.

Para una mejor comprensión del proceso que se sigue en cada modelo, presentamos un ejemplo de cada uno, deteniéndonos especialmente en el modelo deductivo-inductivo por ser el más utilizado en el ámbito escolar.

6.2.1. *Proceso deductivo*

Para su desarrollo seguiremos el ejemplo propuesto por Anguera (1991: 126-127).

En una investigación realizada en el ámbito de la actividad exploratoria en preescolares (Martínez Criado, 1985), y después de un análisis detallado de los planeamientos de diferentes autores de indudable prestigio (Barker, 1930; Weick, 1969; Blunton-Jones, 1972; Smith y Connolly, 1972; Coll, 1978), se distinguieron, al inicio del proceso de categorización, cuatro grandes bloques:

- a) Inactividad y actividades previas o de intermedio (no figuran actividades de contenido temático, aunque sí es posible que se trate de acciones encaminadas a la puesta en marcha de un tema o una actividad).
- b) Actividades desarrolladas en contacto con los materiales y/o instalaciones (se requiere existencia real de contacto y carácter no intencional o fortuito).
- c) Actividades de contacto con adultos y/o niños (que se diferencia del anterior por el «contacto, entendido desde el significado de estar junto a otro», contacto visual, etc.).
- d) Verbalizaciones (todo tipo de manifestaciones verbales con o sin intención comunicativa).

Cada uno de estos cuatro bloques constituye una dimensión o macrocategoría, por lo que cada uno de ellos debe ser objeto de estudio específico con el fin de dar lugar a las correspondientes categorías.

Siguiendo con el ejemplo, se transcriben las correspondientes al bloque A:

- A) Inactividad y actividades previas o de intermedio.
Abarca las acciones momentáneas que apartan al niño del discurso lógico de los diferentes actos en el marco de una actividad concreta. Así, el niño que está dibujando va al lugar donde tiene colgado el abrigo, coge el pañuelo, y se incorpora de nuevo a su actividad de dibujar.
- Está compuesta de las siguientes categorías:
- A1) Inactividad y aislamiento.
- A2) Se desplaza sin objetivo o intención aparente.
- A3) Va directo a zonas ocupadas por niños, instalaciones, material. Entra en relación inmediata.
- A4) Selección o busca algo o alguien concreto de forma activa.
- A5) Participa en la organización de una actividad.
- A6) Actividades sociales y/o de satisfacción de necesidades físicas.

Una vez establecidas las categorías, el autor estableció los límites de cada una, marcando los diferentes rasgos que podrían ser incluidos en cada una de ellas. Por ejemplo, al definir la categoría A1 indica:

«La categoría A1 incluye el caso del niño que solo y distante parece distraído y absorto, permaneciendo en su sitio aunque esté ocupado en su propia automanipulación» (Martínez Criado, 1985: 428)

Mediante una estrategia similar, el autor llegó a establecer, delimitar, definir y precisar las categorías correspondientes a los bloques B), C) y D).

En total, se elaboró un sistema de 39 categorías, desde la que se ha denominado perspectiva deductiva.

6.2.2. *Proceso inductivo*

Si a partir de un registro narrativo o con los medios técnicos que disponemos: magnetófono, vídeo, etc., realizamos adecuados registros, se puede proceder de manera inversa al ejemplo comentado anteriormente. Anguera (1991: 130), especifica paso a paso el proceso seguido, a partir de un registro narrativo en el que a grandes rasgos comprende:

- Fase 1: Registro narrativo.
- Fase 2: Registro semisistemizado.
- Fase 3: Lista total de rasgos.
- Fase 4: Clasificación de rasgos según niveles de respuesta.
- Fase 5: Agrupación de rasgos homogéneos.
- Fase 6: Revisión de la lista de rasgos y elaboración de un sistema provisional de categorías.
- Fase 7: Comprobación del sistema.

6.2.3. *Proceso deductivo-inductivo*

En este tercer ejemplo se parte de un marco teórico para realizar las conceptualizaciones de las macrocategorías y la definición de los límites de cada una de ellas. Posteriormente, a partir de observaciones exploratorias, se realizan las listas de rasgos que se ^{van} fueron adscribiendo a cada categoría en función de las conceptualizaciones y los límites establecidos.

En el siguiente ejemplo se comenzó analizando el diseño curricular de Educación Primaria y el planteamiento que sobre el tema tienen diferentes autores (Coll, 1991, 1992; Slavin, 1983; Vygosky, 1978; Wood, 1980). A partir de estas lecturas se elaboraron las siguientes categorías:

- 1.º El maestro constituye un modelo adulto significativo para su grupo de alumnos.
- 2.º El maestro tiene una importante función en la formación en clase de un grupo humano cohesionado.
- 3.º El maestro tiene un papel esencial como mediador de aprendizajes significativos para los alumnos.

A) Definición de las categorías**a) El maestro constituye un modelo adulto significativo para su grupo de alumnos.**

No podemos considerar al maestro como un individuo más del grupo y en igualdad con los restantes miembros. Su intervención, sus actitudes, su emotividad... conforman para los alumnos un modelo significativo que, junto con el de sus padres y otros adultos relevantes en su vida, contribuyen a forjar una imagen adulta que en buena medida va a incidir en su desarrollo.

El maestro, desde esta perspectiva, deberá desarrollar una serie de actitudes tales como: actitud afectuosa, actitud no autoritaria, mantener conductas coherentes, actitud de respeto y confianza en el niño, etc.

b) El maestro tiene una importante función en cuanto a la formación en la clase de un grupo humano cohesionado.

Uno de los objetivos de la educación es potenciar el desarrollo del niño en su dimensión social, en este sentido la escuela ofrece un marco inigualable en donde este desarrollo puede favorecerse. En la clase debe establecerse una red de comunicación multidireccional entre profesor-grupo de alumnos, el grupo de niños entre sí, el profesor-alumno, profesor-pequeño grupo, alumno-alumno, etc.

Esta red de comunicación ha de hacerse posible a través de distintos recursos, entre los que podemos destacar: favorecer la comunicación informal, creación de distintos espacios que estimulen diversas formas de agrupamiento en contextos variados, la asamblea de clase, etc.

c) El maestro tiene un papel esencial como mediador de aprendizajes significativos para los alumnos.

En la medida que el profesor conoce los procesos de adquisición de conocimiento y las características evolutivas de sus alumnos, diseña su intervención didáctica para promover la construcción de aprendizajes significativos, planteando problemas, diálogos, estableciendo conexiones entre los conocimientos nuevos y aquellos que ya poseía, dedicando especial atención en el discurso a los aspectos que no están claros, que son incongruentes o sobre los que no existe acuerdo en clase, etc.

Una vez analizadas y descritas cada categoría se elaboraron las listas de rasgos, extraídos de la observación directa de los profesores en sus aulas.

La observación se realizó durante una semana, diariamente 45 minutos.

Del análisis de estas cintas se obtuvo un catálogo de estrategias, actitudes y mecanismos de influencia que utilizan los profesores de primero de Educación Primaria.

Una vez extraído este primer listado, se procedió a la adscripción de los rasgos a categoría correspondiente.

Por ejemplo, según el análisis de los documentos, el profesor debe ser «mediador de aprendizajes significativos». Textualmente dice:

«El profesor ocupa un lugar importante en la interacción educativa. En la medida que conoce los procesos de adquisición de conocimiento y las características

evolutivas de sus alumnos, diseña su intervención didáctica para promover la construcción de aprendizajes significativos..., plantea problemas, orienta la conversación, etc.» (Diseño Curricular de Primaria: 104).

Con este texto de referencia, todas las observaciones hechas sobre el profesor, planteando problemas, orientando el diálogo, ofreciendo información, etc., entran a formar parte de esta categoría.

Así se continuó adscribiendo cada uno de los rasgos a las tres categorías comentadas.

Se observaron comportamientos y actitudes en los profesores cuya presencia era negativa. Estas conductas también han sido incluidas en la categoría, pero con formulación positiva. Por ejemplo, si hemos observado una actitud racista, como rasgo en la categoría (modelo adulto significativo), se ha planteado como «no racista ni sexista», o si se observa que impide la participación del alumno cuando el profesor está explicando un concepto, se ha formulado como: «anima a intervenir».

La adscripción de los rasgos a las diferentes categorías fueron realizadas por dos observadores, los mismos que previamente habían visualizado las cintas y obtenido las listas de rasgos.

Las categorías fueron las siguientes:

1. Servir de modelo adulto significativo.
 - 1.1. Respeto y confianza en el niño.
 - 1.2. Seguridad en sí mismo.
 - 1.3. Conducta no autoritaria.
 - 1.4. Manifiesta opiniones que sirven para combatir estereotipos sociales.
2. Crear un grupo humano cohesionado.
 - 2.1. Establece una red de comunicación multidireccional.
 - 2.2. Favorece la comunicación informal.
 - 2.3. Crea un clima afectivo positivo.
 - 2.4. Planifica e interviene en el trabajo grupal de clase.
 - 2.5. Establece relación con distintas personas que entren en el aula.
 - 2.6. Establece relaciones entre los grupos.
3. Mediador de aprendizajes significativos.
 - 3.1. Plantea problemas.
 - 3.2. Establece conexiones entre el conocimiento nuevo y el que el niño ya tenía.
 - 3.3. Ofrece información relevante.
 - 3.4. Observa el proceso y se detiene en aquellos aspectos que parecen no estar claros.
 - 3.5. Crea situaciones que permiten al alumno reflexionar sobre su propio proceso de aprendizaje.
 - 3.6. Organiza debates.
 - 3.7. Orienta al diálogo.
 - 3.8. Ayuda a los que lo necesitan a expresar sus ideas.
 - 3.9. Anima a intervenir.

Cada uno de estos rasgos fueron conceptualizados y definidos sus límites. Por ejemplo: *Respeto y confianza en el niño*. Se entiende por este rasgo la conducta del profesor que sin permitir faltas de respeto fomenta un clima de confianza, no ridiculizando a nadie y animando a todos a expresar sus ideas. Incluye conductas de ánimo hacia las tareas que realizan los alumnos, de invitación a intervenir en clase, de aceptación de sugerencias. Se diferencia de la 1.2, en que ésta va referida a la mediación del profesor cuando el alumno no interviene por timidez o falta de seguridad en sí mismo.

B) Planilla de registro

Los rasgos de cada categoría se registran cada vez que aparecen en un intervalo de tiempo determinado.

Para realizarlo es necesaria una planilla donde se contemplen las apariciones de la conducta a observar y todos los indicadores que interesen al estudio. Para este caso se confeccionó la siguiente:

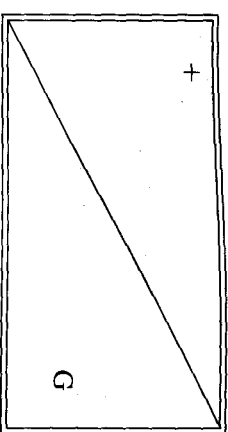
El profesor plantea problemas																			
Ofrece información																			
Organiza debates																			
Anima a intervenir																			
Orienta el diálogo																			
Etcétera																			

Cada cuadrícula equivale a un período de observación de 5 minutos. El procedimiento seguido en el registro fue:

- (+) Presencia de la conducta categorizada.
- (−) Conducta contraria.
- (A) Cuando el maestro se dirige a la clase en general.
- (G) Cuando el maestro se dirige a algún grupo de trabajo cooperativo en particular.
- (I) Cuando el maestro se dirige a un alumno en particular.

El observador registrará la presencia o ausencia del rasgo y si va dirigida a un individuo, a un grupo, o a la clase en su conjunto:

Ejemplo: En la categoría 3 podemos encontrar el siguiente registro correspondiente al rasgo 3.9, lo que significa que el maestro ha estado animando a intervenir a los alumnos que trabajaban en grupo.



Cada cinco minutos el observador anota las conductas que aparecen en un cuadro distinto. Al final de la sesión se contabiliza la frecuencia de dicha conducta a lo largo de los 45 minutos de observación.

6.3. Fiabilidad de las observaciones

Una confusión general, en lo referente a la fiabilidad de las observaciones, es la identificación de este concepto con los coeficientes de concordancia de los observadores.

La concordancia se define como la consistencia entre observadores cuando éstos están simultáneamente codificando los mismos eventos en una clase (concordancia interobservadores), o un observador consigo mismo en distintos momentos (concordancia intraobservadores). Hay un tercer tipo de concordancia que actualmente ocupa un lugar primordial en la metodología observacional; es la concordancia consensuada. Cuando la conducta ha sido grabada en vídeo o, si sólo interesa la conducta verbal, en magnetófono, los observadores deben lograr un acuerdo previo sobre la asignación de cada código a su correspondiente unidad de conducta. Si se logra este acuerdo antes del registro, se habrá conseguido no sólo un registro único sino además que el sistema de categorías quede perfectamente delimitado y definido.

La concordancia es pues lo primero, pero no lo fundamental, en la interpretación de los resultados de los estudios observacionales, y menos, sinónimo de fiabilidad como se asume con frecuencia. Así, por ejemplo, los observadores pueden recolectar datos muy poco fiables, si la varianza entre sujetos es pequeña en relación a la varianza en los sujetos, independientemente del grado de concordancia observacional.

¿Cuándo debe ser medida la concordancia? En orden a minimizar la posibilidad de que los observadores sean la principal fuente de la baja fiabilidad de los datos observacionales, el investigador debe asegurarse de que estén adecuadamente entrenados antes de proceder a la recogida de datos. Sin embargo, aunque los observadores hayan sido suficientemente preparados, el grado de acuerdo puede deteriorarse a lo largo del tiempo de observación, por lo que es recomendable su comprobación periódica hasta el final del estudio.

La fiabilidad, aun siendo un término restrictivo desde el punto de vista de la psicometría (aprecia cuán precisa es una medida o cuánto se acerca a la «verdad»), en la metodología observacional es un término de mayor amplitud al ser tratado de diferentes formas en diferentes situaciones.

Podemos concretar tres formas diferentes de entenderla (Medley y Mizel, 1963; Mitchell, 1978; Frick y Semmel, 1978):

- Coefficiente de concordancia. Se refiere a las observaciones realizadas por diferentes observadores en un mismo momento.
- Concepto de fiabilidad, desde el punto de vista de la teoría psicométrica, obtenida a través del coeficiente de correlación.
- Teoría de la generalizabilidad. Se trata de una teoría multivariada que tiene en cuenta diferentes fuentes de error, incluyendo las diferencias individuales entre los sujetos entrevistados. A través del análisis de varianzas se puede cuantificar la importancia de cada fuente de variación.

6.3.1. Índices de concordancia

Hay diferentes formas de realizar los índices de concordancia:

A) Porcentaje simple de acuerdo

Puede ser calculado de dos formas: la primera consiste en comparar en cada categoría el acuerdo de los eventos individuales (concordancia nominal). La segunda, comparando la frecuencia total de las categorías a lo largo de un número de eventos (concordancia marginal).

Ejemplo: En la siguiente tabla de contingencia se recogen las frecuencias de aparición de eventos en cada categoría, para el observador 1 y el observador 2.

	Observador 2					n_{i+}
	C1	C2	C3	C4	C5	
C1	2	2	0	0	0	4
C2	0	6	0	0	0	6
C3	0	0	4	0	0	4
C4	0	1	4	5	0	10
C5	0	3	0	0	3	6
n_{+i}	2	12	8	5	3	30

El acuerdo nominal para C categorías es definido por:

$$P_o = \frac{1}{N} \sum_{i=1}^C n_{ii}$$

El acuerdo nominal para las cinco categorías es:

$$P_o = \frac{1}{30} (2 + 6 + 4 + 5 + 3) = \frac{20}{30} = 0,67$$

El segundo porcentaje de acuerdo simple es el acuerdo marginal. Este índice se utiliza cuando las categorías marginales (frecuencias totales) están disponibles.

$$P_{oi} = \frac{\min(n_{i+}, n_{+i})}{\max(n_{i+}, n_{+i})}$$

n_{i+} y n_{+i} son respectivamente observaciones marginales para cada una de las categorías. mín y máx indican que la marginal menor se divide entre la marginal mayor.

El índice de acuerdo marginal total es:

$$P_o = \frac{1}{C} \sum_{i=1}^C P_{oi}$$

Ejemplo: El porcentaje de acuerdo marginal en las cinco categorías es:

$$P_o = \frac{1}{5} \left(\frac{2}{4} + \frac{6}{12} - \frac{4}{8} + \frac{5}{10} + \frac{3}{6} + \dots \right) = 0,50$$

La concordancia nominal tiende a ser más rígida que la concordancia marginal. Éste puede ser apropiado si los datos van a ser analizados sobre la base del total de categorías de frecuencias (Frick y Semmel, 1978).

B) Otros índices de concordancia: el índice de Scott y el coeficiente de Kappa

B.1) El índice de Scott

Scott (1955) argumenta que las medidas de porcentaje simple de concordancia pueden estar influidas por la concordancia al azar. Por esto propone el coeficiente π , que

estima el grado en que la concordancia al azar ha sido superada, cuando son comparadas las puntuaciones de dos observadores.

$$\pi = \frac{P_o - P_e}{1 - P_e}$$

$P_o = \frac{1}{N} \sum_{i=1}^C n_{ii}$ = a concordancia nominal para C categorías ($C \geq 2$).

$P_e = \sum_{i=1}^C P_i^2$ = a concordancia al azar para C categorías.

B.2) Coeficiente Kappa

Fue propuesto por Cohen (1960). En este coeficiente el acuerdo por azar, P_e , está basado en las distribuciones marginales observadas más que en las distribuciones esperadas; y en él no se requiere el supuesto de que los marginales sean simétricos y proporcionales a la población de marginales conocidos. La concordancia nominal, P_o , para C categorías ($C \geq 2$), es calculada de la misma forma que en el coeficiente de Scott, pero P_e es definido por:

$$K = \frac{P_o - P_e}{1 - P_e}$$

$$P_e = \frac{1}{N^2} \sum_{i=1}^C (n_{i+})(n_{+i})$$

donde n_{i+} y n_{+i} son respectivamente observaciones marginales por cada categoría. En la tabla de contingencia de 5×5 expuesta anteriormente:

$$P_e = \frac{1}{30^2} (4 \times 2 + 12 \times 6 + 8 \times 4 + 10 \times 5 + 6 \times 3) = \frac{180}{900} = 0,20$$

$$P_o = \frac{1}{30} (2 + 6 + 4 + 5 + 3) = \frac{20}{30} = 0,67$$

$$K = \frac{0,67 - 0,20}{1 - 0,20} = 0,59$$

Ambos estadísticos se pueden utilizar siempre que:

1. Las categorías sean nominales y mutuamente excluyentes.
2. Las categorías han de ser independientes.
3. Los observadores han de actuar también independientemente.

6.3.2. Concepto de fiabilidad a través de coeficientes de correlación

Si el estudio está bien diseñado, el grado en que la no concordancia limita la fiabilidad de las medidas observacionales puede ser determinado por la correlación intraclass. Una correlación intraclass es una estimación de la razón entre la verdadera varianza de las conductas (varianza entre los sujetos) y la varianza obtenida. La varianza obtenida incluye la verdadera varianza más la varianza causada por el error de medida, y algún otro factor incluido en el error de varianza.

El principal problema encontrado con este índice es que la cantidad de varianza verdadera fluctúa cuando, o si, las situaciones seleccionadas para ser registradas son altamente diferentes o muy similares.

El coeficiente de correlación más utilizado es el producto-momento de Pearson, basado en los pares de puntuaciones obtenidos en las siguientes situaciones:

- Dos observadores que han observado simultáneamente una sesión (fiabilidad interobservadores).
- Un observador que registra la sesión dos veces en dos períodos de tiempo distintos (fiabilidad intraobservador).
- Dos observadores registrando una misma sesión de observación pero en períodos de tiempo distintos (fiabilidad como homogeneidad).
- Dos observadores registrando dos sesiones de observación similares (fiabilidad como equivalencia).
- Un observador acerca de dos sesiones de observación, de dos períodos de tiempo diferentes (fiabilidad como constancia).

Los valores de estos índices oscilan entre -1 a $+1$.

Los coeficientes de correlación deben basarse en un número suficiente de datos, como mínimo 8 o 10 pares de puntuaciones (Hartmann, 1979). Como indicadores de la fiabilidad, sólo supone la existencia de medidas absolutas idénticas entre los observadores en los diferentes momentos de observación, independientemente de la posición relativa que tengan las medidas obtenidas en los diferentes momentos de observación de cada observador (Silva, 1987).

$$r = \frac{\sum x' \cdot y'}{N \sigma_x \sigma_y}$$

x' = observaciones de $X_i = (x_i - \bar{x})$.

y' = observaciones de $Y_i = (y_i - \bar{y})$.

σ_x = desviación típica de la variable x .

σ_y = desviación típica de la variable y .

Cuando los datos se presentan dicotomizados se puede utilizar el coeficiente φ :

$$\varphi = \frac{(AD) - (BC)}{\sqrt{(A+B)(C+D)(A+C)(B+D)}}$$

El coeficiente de fiabilidad, calculado por la fórmula de la correlación producto-momento, tiene una interpretación matemática precisa, r es igual a la proporción de varianza de la puntuación total, no debida al error, y el grado de asociación lineal entre los datos de dos observadores; r^2 es igual a la proporción de varianza de la puntuación de un observador que es predecible por el conocimiento de las puntuaciones de otro observador.

Para conocer la significación del coeficiente de correlación se calcula:

$$t = (\sqrt{N-2}) \left(\frac{r}{\sqrt{1-r^2}} \right)$$

y se consulta la tabla t con $N - 2$ grados de libertad.

6.3.3. Teoría de la generalizabilidad

Fue desarrollada por Cronbach, Gleser, Nanda y Rajaratham para unificar los diversos métodos anteriores de estimación del error, gracias al concepto estadístico de muestreo de facetas múltiples. Es decir, además de las diferencias individuales, hay otras fuentes de variación (personalidad del observador, características del sujeto observado, características del sistema de registro, etc.) que afectan a los resultados. Cada una de las características puede constituir una faceta de un plan de observación sistemática. Aplicando las técnicas de análisis de varianza a estas observaciones se puede cuantificar la importancia de cada fuente de variación. La puntuación verdadera se puede definir como: «La esperanza matemática de todas las observaciones posibles» y el error como «una fluctuación muestral correspondiente a la extracción aleatoria de ciertos niveles de las facetas consideradas» (elección de determinados momentos, determinados observadores, ciertas operaciones, etc.) (Brennan y Kane, 1979).

La teoría estadística nos dirá en qué intervalo se sitúa la puntuación verdadera cuando se emplean ciertas formas de muestreo y qué progresos conseguiremos si elegimos de otra forma las muestras.

La teoría de la generalizabilidad nos ofrece así un marco más satisfactorio para estimar la fiabilidad y los márgenes de error, ya que es suficientemente globalizado para como para adaptarse a las condiciones particulares de cada objeto de medida (Touner y Cardinet, 1979).

A) Conceptos de la teoría de la generalizabilidad

Los dos primeros conceptos a aclarar son los de Población y Universo. El término Población hace referencia a los objetos de medida; mientras que el término Universo denota las condiciones de evaluación de los objetos.

Cuando hablamos de puntuación universo, atendemos a dos factores:

- Puntuación universo para el sujeto «P» (μ_P): una persona «P» es evaluada a través de todas las categorías posibles (i), y se calcula la media de todos los (π) obtenidos.
- Puntuación universo ligada a la categoría «i» para: todas las personas de la población (μ_i): «Todos los sujetos de una misma población son evaluados en una misma categoría (i), obteniendo una media».

En general, la puntuación universo explica el hecho de que lo que interpreta una medida es la estimación, a partir de una muestra de datos observados, de un valor teórico inobservable. Se intenta conocer la media de todos los valores que se obtendrían si se efectuasen las observaciones en todas las condiciones posibles.

- El término Faceta designa cada una de las características de la situación que es susceptible de ser modificada de una observación a otra, pudiendo hacer que varíe el resultado.

- A cada una de las manifestaciones posibles de una faceta (cada observador, cada sujeto, cada instrumento) es designado nivel de la faceta. El observador número 2 es designado como un nivel de la faceta observadores.

Como ya hemos señalado, la puntuación universo de una persona «P», que es el dato ideal, representa la media de las puntuaciones de dicha persona calculada sobre todas las observaciones posibles. El observador puede utilizar, o bien la puntuación observada, o bien una función de dicha puntuación, para poder estimar el valor de la puntuación universo. Así se generalizará de la muestra a la población. De aquí que el problema de fiabilidad es el problema de la precisión de la generalización. La generalizabilidad es el grado por el cual podemos generalizar un resultado obtenido en unas condiciones particulares, a un valor teórico buscado.

- El coeficiente de generalizabilidad trata de estimar en qué medida se puede generalizar a partir de la media observada en esas condiciones, a la media de todas las observaciones posibles. Existen diferentes coeficientes de generalizabilidad, en función del plan de observación, según el objeto de medida y según los universos de generalización elegidos.
- El universo de generalización es el conjunto de condiciones a las que se quieren generalizar los resultados observados en esas condiciones particulares. Esto resulta de elegir un subconjunto de condiciones «admisibles» en el conjunto original de todas las condiciones posibles.

Todo evaluador tiene en mente un universo al que quiere generalizar sus observaciones. Este universo define las fuentes de variación que le interesan y que tiene que tener en cuenta.

Por tanto debe estimar todos los componentes de varianza de las observaciones, a través de un estudio previo, llamado estudio de generalizabilidad (G). Este estudio implica calcular la parte de la varianza total que es atribuible a las facetas y a sus interacciones. Una vez realizado el estudio G, definimos un nuevo plan de observación que tratará de reducir los componentes no deseados de la varianza de las puntuaciones. Éste es el objeto de un estudio de decisión, estudio D, partiendo de la información del estudio G (Buendía, 1990, 1992).

B) Fases de la teoría de la generalizabilidad

Las fases las vamos a presentar siguiendo el orden de realización y por tanto de importación sucesiva de información (Tournier y Cardinet, 1979). Es evidente que los problemas que plantea cada fase deben quedar previstos al establecer el plan de investigación.

La primera fase es la de identificación y de organización de los datos. Se eligen las facetas a tener en cuenta (y por tanto, también aquellas que permanecerán ocultas, o sea, aquellas cuyo efecto no podrá controlarse). Se precisan las interrelaciones entre las facetas estudiadas. Se decide el número de niveles muestreados en cada faceta. Con esta información se definirá el *plan de observación*, y a través de un análisis de varianza se calculan los cuadrados medios de cada fuente de variación del plan definido. La segunda fase proporciona información sobre el modo de muestreo. Se distinguen, básicamente, dos tipos de muestreo: muestreo aleatorio y muestreo exhaustivo. El primero puede llevarse a cabo en un conjunto de niveles finitos o infinitos, el segundo sólo puede realizarse cuando el número de niveles o de casos sea finito. Como consecuencia podemos encontrarnos con tres modelos diferentes: puramente aleatorio, aleatorio finito, y fijo. Estas informaciones permiten calcular los componentes de varianza (según el modelo mixto) para cada faceta del plan de observación. Definen el *plan de estimación*.

La tercera fase comienza cuando se precisa el papel de cada faceta en la medida. Las facetas, que son objeto de estudio, forman los aspectos de diferenciación. Las facetas de condiciones de observación, o sea, las referidas a los instrumentos de medida, forman los aspectos de instrumentación. Estas informaciones definen el *plan de medida*. Ello permite calcular la varianza de diferenciación, o varianza vertical, y la varianza de error.

La cuarta fase conduce a la elección de un nuevo plan, enfocado a la toma de decisiones (estudios D de Cronbach). Este nuevo plan, llamado *plan de optimización* por Tournier y Cardinet, se establece en base a la información obtenida en la fase precedente, cambiando el plan de observación, el de estimación, el de medición o bien varios de éstos a la vez. El coeficiente de generalizabilidad final deberá controlar que se ha obtenido una suficiente fiabilidad.

ANÁLISIS DE LOS DATOS OBTENIDOS MEDIANTE OBSERVACIÓN

La observación proporciona gran cantidad de datos que son necesarios analizar para extraer de ellos la mayor información posible, en función de los objetivos de la investigación y del tipo de medida empleada.

Como hemos visto en los diseños de observación, se pueden recoger cuatro tipos de informaciones:

- Frecuencia de la conducta.
- Orden de aparición.
- Duración.
- Intensidad.

Quera (1993) establece dos categorías de análisis: macroanálisis y microanálisis.

En el macroanálisis los datos pueden analizarse globalmente, con tablas de frecuencias, duraciones medias, etc.

Por ejemplo, diferentes sesiones de observación de individuos que intentan controlar la adicción al tabaco, realizando una actividad sustitutiva (por ejemplo, ocupar las manos con un pañillo).

El macroanálisis puede realizarse tanto en datos obtenidos con diseños secuenciales como con datos obtenidos en diseños transversales.

Por ejemplo, si deseamos conocer la relación entre conductas autoritarias de los padres e introversión de los hijos, habrá que observar durante diferentes sesiones a varias parejas de padres en el hogar y posteriormente a los hijos en el colegio, o con los amigos, para poder establecer la relación a lo largo de los años, o el tiempo que se determine (análisis longitudinal).

Si comparamos las medidas globales obtenidas simultáneamente, por ejemplo, relación autoritaria de diferentes padres hacia hijos de diferentes edades, se realizará un análisis transversal.

Con estos datos se pueden realizar análisis estadísticos univariados, bivariados o multivariados (valores medios de las variables, análisis de varianza o covarianza, análisis factorial etc.).

El microanálisis se realiza cuando el objetivo de la investigación es conocer la aparición o no de determinados rasgos de conducta y sus respectivas relaciones, bien entre ellos o con otras variables de interés, en cada sesión de observación.

Por ejemplo, se realiza un microanálisis de la conducta interactiva de dos alumnos en el aula, cuando analizamos en cada sesión de observación la secuencia que siguen las diferentes conductas:

- Pide ayuda.
- Facilita información.
- Plantea dudas.
- Facilita ayuda, etc.

El registro se habrá realizado con la frecuencia y orden de aparición de la conducta, teniendo en cuenta el momento de ocurrencia o el intervalo de duración.

El análisis se realiza sobre los siguientes tipos de secuencias (Bakeman y Quera, 1996):

A) De eventos

Las unidades de conducta son representadas por una serie de códigos, mutuamente excluyentes. Obtenemos información de la frecuencia y el orden en el que aparecen conductas fugaces o estados de conducta, en los que se prescinde del tiempo tanto de aparición como de duración.

Ejemplo de registro:

- | | |
|------------|--------------|
| 1.ª sesión | acbadbcbad. |
| 2.ª sesión | cbacbabcbad. |

B) De estados

Cuando la observación se realiza sobre una conducta duradera e interesa registrar bien el momento de aparición o el intervalo de duración. Conductas: A B C D.

Ejemplo de registro:

1.ª sesión A (9:10'), B (9:16'), D (9:30'), C (9:31'), etc.

Indica que la conducta A se inició a las 9 horas 10 minutos y duró hasta las 9 horas 16 minutos que se inició la conducta B. Ésta termina a las 9 horas 30 minutos que aparece la siguiente, y así sucesivamente.

C) De eventos y estados

Son las secuencias mas complejas, y se registran las conductas de eventos y las conductas de estado que aparecen simultáneamente.

Ejemplo de registro:

Eventos: a b c d.

Estados: A B C D.

1.ª sesión A (13:10'; c); C (13:17'; abd); D (13:19'; cd).

Indica que la conducta A que es una conducta de estado se inicia a las 13 horas 10 minutos y dura hasta las 13 horas y 17 minutos que se inicia la conducta C. A la vez se ha producido el evento c. Y así sucesivamente.

D) Secuencias de intervalos

Son conjuntos de conductas, no necesariamente mutuamente excluyentes, aparecidas en intervalos de tiempo constantes.

En cada intervalo puede aparecer desde cero códigos hasta todos los posibles, en función de las apariciones de conducta en ese intervalo de tiempo.

Ejemplo de registro:

Conductas	5'	5'	5'	Totales
A	+ (1.º)		+ (2.º)	2
B	+ (2.º)	+ (1.º)	+ (3.º)	3
C			+ (1.º)	1

En el primer intervalo ha aparecido la conducta ab. En el segundo sólo la conducta b. Y en el tercero ABC.

Bakeman (1968) estableció una clasificación de los datos secuenciales en función de dos dimensiones bipolares: que las unidades de conducta sean concurrentes o no, y la consideración temporal.

El cruce de estas dos dimensiones dio lugar a cuatro tipos de datos:

- Datos tipo 1. Las unidades de conducta son eventos (o estados en los que no se contabiliza el tiempo) en los que se registra la frecuencia de aparición y el orden. Son unidades mutuamente excluyentes. Ejemplo: ríe, llora, canta, habla. Son equivalentes a las secuencias de eventos.
- Datos tipo 2. Son datos concurrentes, en los que no se registra el tiempo. Implica registrar varias unidades de conducta simultáneamente, de una misma lista de rasgos o de dos sistemas mutuamente excluyentes (ríe, está sentado, mueve las manos) (habla, balancea la cabeza, camina), etc.
- Datos tipo 3. Son datos semejantes al tipo 1, pero incluyen la dimensión temporal. Ejemplo: ríe, 1'; llora, 2'; habla, 5'; etc.
- Datos tipo 4. Son semejantes a los tipo 2, pero incluye el tiempo de duración de las unidades de conductas observadas. Son pues unidades de conducta, que ocurren simultáneamente y en cuyo registro también se recoge el tiempo de duración de la coocurrencia.

Las técnicas de análisis más utilizadas para datos secuenciales son: la técnica de retardos (Sackett, 1978); los modelos de Markov, y las series temporales.

Para un análisis detallado recomendamos los trabajos de Bakeman (1978); Anguera (1990); Quera (1993).

En la actualidad existen importantes programas de ordenador, para el análisis de datos secuenciales. Son especialmente interesantes el SDIS y GSEQ (Bakeman y Quera, 1996).

Cuando los datos han sido obtenidos con diseños transversales, las observaciones se presentan en tablas de contingencia, por lo que los análisis estaban casi reducidos a la aplicación de chi-cuadrado. Hoy, gracias a la aplicación de los modelos lineales logarítmicos «log-linear» (cuando sólo interesa la presencia o ausencia de asociación entre las variables, sin direccionalidad u orden causal entre ellas) y el análisis «logit» (cuando una variable asume el estatus de explicativa o independiente, y otra el de respuesta o dependiente), se han ampliado enormemente las posibilidades de análisis y el establecimiento de modelos de mayor complejidad (Anguera, 1990: 208).

8. EL INFORME DE LA INVESTIGACIÓN OBSERVACIONAL

El informe final de la investigación realizada con la metodología observacional no difiere en gran medida de los modelos planteados anteriormente, por ejemplo en la investigación por encuesta, página 151, o el informe en el proceso de la investigación, páginas 42-53. Sin embargo, es necesario incluir algunos aspectos que aquí son imprescindibles para la réplica de la investigación, o, simplemente, para una correcta divulgación de la misma. Los apartados que deben incluirse son los mostrados en el gráfico de la página siguiente.